

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
DEPARTAMENTO ACADÊMICO DE INFORMÁTICA  
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

**ANDRESSA OLIVEIRA SOUZA**

**MAPEAMENTO ONTOLÓGICO PARA CLASSIFICAR MENSAGENS  
SIGNIFICATIVAS EM ATIVIDADES COLABORATIVAS**

**TRABALHO DE CONCLUSÃO DE CURSO**

**PONTA GROSSA**

**2018**

**ANDRESSA OLIVEIRA SOUZA**

**MAPEAMENTO ONTOLÓGICO PARA CLASSIFICAR MENSAGENS  
SIGNIFICATIVAS EM ATIVIDADES COLABORATIVAS**

Trabalho de Conclusão de Curso apresentado como requisito parcial à obtenção do título de Bacharel em Ciência da Computação, do Departamento Acadêmico de Informática, da Universidade Tecnológica Federal do Paraná.

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Simone Nasser Matos

**PONTA GROSSA**

**2018**



Ministério da Educação  
**Universidade Tecnológica Federal do Paraná**  
Câmpus Ponta Grossa

Diretoria de Graduação e Educação Profissional  
Departamento Acadêmico de Informática  
Bacharelado em Ciência da Computação



---

## TERMO DE APROVAÇÃO

### MAPEAMENTO ONTOLÓGICO PARA CLASSIFICAR MENSAGENS SIGNIFICATIVAS EM ATIVIDADES COLABORATIVAS

por

ANDRESSA OLIVEIRA SOUZA

Este Trabalho de Conclusão de Curso (TCC) foi apresentado em 20 de novembro de 2018 como requisito parcial para a obtenção do título de Bacharel em Ciência da Computação. A candidata foi arguida pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho aprovado.

---

Prof<sup>a</sup>. Dra. Simone Nasser Matos  
Orientadora

---

Prof. MSc Clayton Kossoski  
Membro titular

---

Prof. MSc Vinicius Camargo Andrade  
Membro titular

---

Prof<sup>a</sup>. Dra. Helyane Borges  
Responsável pelo Trabalho de Conclusão  
de Curso

---

Prof. Dr. Saulo Jorge Beltrão de Queiroz  
Coordenador do Curso

- O Termo de Aprovação assinado encontra-se na Coordenação do Curso -

Dedico este trabalho à minha família, em especial minha mãe que sempre acreditou em mim e ao Matheus que virou estrela no céu antes da hora.

## **AGRADECIMENTOS**

Primeiramente, gostaria de agradecer a minha família e em especial a minha mãe por todo o suporte e apoio para que eu pudesse realizar este sonho. Sem minha família seria impossível chegar onde hoje estou. Que um dia eu me torne metade da mulher que minha mãe é.

Agradeço minha orientadora Prof<sup>ª</sup>. Dra. Simone Nasser Matos, por ter me aceitado como aluna de iniciação científica e sempre estar ao meu lado me guiando e transcendendo sabedoria em todas as aulas e reuniões nessa árdua trajetória.

Aos meus amigos de condomínio que sempre estiveram ao meu lado em momentos de alegria, ansiedade e tristeza. Amigos estes que se disponibilizaram a ler e assistir inúmeras vezes etapas deste trabalho. Agradeço pelos cafés da tarde em dias ensolarados no corredor e em especial meus veteranos Milton e Elder que sempre se dispuseram a compartilhar conhecimento e sanar minhas dúvidas.

Agradeço também todos os meus professores do DAINF que me ensinaram simplesmente tudo que precisava para chegar onde estou, em especial alguns nomes como a Prof<sup>ª</sup>. Dra Sheila Moraes de Almeida, Prof. Dr. Erikson Freitas de Moraes que talvez não saibam, mas em momentos de dificuldade disseram coisas que eu precisava ouvir para não desistir e acreditar mais em mim mesma.

Um abraço apertado aos meus amigos sobreviventes de sala que compartilharam trabalhos, momentos e lágrimas durante o curso.

Para finalizar, gostaria de deixar registrado meu reconhecimento a tudo e todos que de alguma forma contribuíram para que eu realizasse este sonho. Foram incontáveis às vezes em que a insegurança bateu em minha porta e eu me mantive de pé e lutei contra todas as pedras no meu caminho.

Às vezes são as pessoas que menos  
imaginamos, que fazem as coisas que  
ninguém pode imaginar.  
(TURING, Alan, 1954)

## RESUMO

SOUZA, A. O. **Mapeamento Ontológico para Classificar Mensagens Significativas de Atividades Colaborativas**. 2018. 78 f. Trabalho de Conclusão de Curso Bacharelado em Ciência da Computação - Universidade Tecnológica Federal do Paraná. Ponta Grossa, 2018.

A colaboração permite que os indivíduos possam compartilhar objetos, informações e ideias na tomada de decisão, mas é difícil mensurar a participação de cada pessoa. As informações em um ambiente colaborativo, em sua maioria, são modeladas em diferentes pontos de vista sobre o conhecimento do contexto. Normalmente, cada sistema possui sua própria linguagem e sintaxe para a representação do conhecimento, formando uma semântica local para cada um deles o que dificulta o processamento de dados. As ontologias ajudam a resolver problemas relacionados a semântica porque são capazes de representar, compartilhar e reutilizar conhecimento em comum. Este trabalho tem a finalidade de realizar o mapeamento de interações entre os indivíduos em um ambiente de atividades colaborativas com o uso de ontologias para avaliar as mensagens significativas. A aplicação da ontologia proposta foi aplicada nas mensagens trocadas entre os participantes do *Collabora*, que é um ambiente de aprendizagem colaborativa para promover a colaboração entre estudantes por meio de um *chat* e permite a troca de mensagens, *links* e arquivos para a resolução de atividades. A ontologia proposta consegue identificar a partir das interações entre os participantes quais mensagens podem ser classificadas como significativas.

**Palavras-chave:** Ontologias. Mapeamento Ontológico. Ambientes Colaborativos.

## ABSTRACT

SOUZA, A. O. **Ontological Mapping to Classify Significant Messages in Collaborative Activities**. 2018. 78 p. Work of Conclusion Course Graduation in Computer Science - Federal Technology University - Paraná. Ponta Grossa, 2018.

Collaboration allows individual to share objects, information and ideas to make decisions, but it is hard to measure the participation. The information in a collaborative environment, for the most part, are in different modeled on different points of view about the knowledge of the context. Normally, each system has its own language and syntax for a representation of knowledge, forming a local semantics for each of them, or making data processing difficult. This work has the purpose of mapping interactions between individuals in a collaborative environment of activities with the use of ontologies to evaluate meaningful messages. The application of the proposed ontology was applied in the exchanged messages between the Collabora participants, which is a collaborative learning environment to promote collaboration among students through a chat and allows the exchange of messages, links and files for the activities resolution. The proposed ontology was able to identify from interactions between the participants which messages could be classified as meaningful.

**Keywords:** Ontologies. Ontology Mapping. Collaborative Environments.



## LISTA DE ILUSTRAÇÕES

Figura 1 - Níveis Ontológicos .....	16
Figura 2 - Exemplo de Alinhamento de Ontologias .....	19
Figura 3 - Processo de Desenvolvimento Ontology Development 101 .....	23
Figura 4 - Processo de Desenvolvimento Mentology .....	25
Figura 5 – Processo de Desenvolvimento On-To-Knowledge .....	26
Figura 6 - Uso de ontologias para formalização de conhecimento em Engenharia de <i>Software</i> .....	30
Figura 7 - Estrutura do Modelo 3C .....	34
Figura 8 - Tela de atividade sendo realizada no <i>Collabora</i> .....	36
Figura 9 - Tela de análise de mensagens no <i>Collabora</i> .....	37
Figura 11 - Exemplo de colaboração.....	41
Figura 12 - Mapa Conceitual de Representação de Parâmetros para Mensagens ...	43
Figura 13 - Criação de Classes da Ontologia.....	44
Figura 14 - Configuração da Propriedade <i>Data Properties</i> .....	45
Figura 15 - Configuração da propriedade <i>Object Properties</i> .....	46
Figura 16 - Exemplo de um Axioma na Ontologia .....	47
Figura 17 - Exemplo de um axioma padrão .....	47
Figura 18 - Exemplo de um axioma para significativa .....	47
Figura 19 - Fluxo de Criação de Base de Conhecimento .....	48
Figura 20 - Exemplo de classificação de mensagem .....	52
Figura 21 - Exemplo 2 de classificação de mensagens significativas .....	53
Quadro 1 - Ontologias de Fundamentação .....	17
Quadro 2 – Comparativo entre ferramentas de ontologias segundo Alatrish .....	21
Quadro 3 - Algumas medidas de avaliação utilizadas no <i>Collabora</i> .....	38
Quadro 4 - Conteúdo de questões do <i>dataset</i> do <i>Collabora</i> .....	50

## LISTA DE ABREVIATURAS

AVA	Ambiente Virtual de Aprendizagem
CSLC	Aprendizagem Colaborativa Apoiada por Computador
DAINF	Departamento Acadêmico de Ciência da Computação
OA	Objeto de Aprendizagem
OTKM	<i>On-To-Knowledge Methodology</i>
PLN	Processamento de Linguagem Natural
UFO	Ontologia de Fundamentação Unificada

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>12</b>
1.1 OBJETIVOS.....	13
1.2 ORGANIZAÇÃO DO TRABALHO.....	13
<b>2 ONTOLOGIAS: UMA VISÃO GERAL</b> .....	<b>14</b>
2.1 CONCEITOS.....	14
2.2 NÍVEIS ONTOLÓGICOS.....	16
2.3 MAPEAMENTO ONTOLÓGICO .....	18
2.4 FERRAMENTAS PARA MAPEAMENTO.....	20
2.4.1 Protégé .....	21
2.5 ESTRATÉGIAS DE MAPEAMENTO.....	22
2.5.1 Ontology Development 101 .....	22
2.5.2 Mentholgy .....	24
2.5.3 On-To-Knowledge.....	26
2.6 CONSIDERAÇÕES DO CAPÍTULO .....	27
<b>3 ESTADO DA ARTE</b> .....	<b>28</b>
3.1 ONTOLOGIA E DESENVOLVIMENTO DE SOFTWARE .....	28
3.2 ONTOLOGIA E MAPEAMENTO .....	30
3.2.1 Trabalhos Relacionados .....	31
3.3 CONSIDERAÇÕES DO CAPÍTULO .....	32
<b>4 ATIVIDADES COLABORATIVAS</b> .....	<b>33</b>
4.1 CONCEITOS GERAIS .....	33
4.2 FERRAMENTAS COLABORATIVAS.....	35
4.3 COLLABORA.....	36
4.4 CONSIDERAÇÕES DO CAPÍTULO .....	38
<b>5 MAPEAMENTO ONTOLÓGICO PROPOSTO PARA CLASSIFICAÇÃO DE MENSAGENS</b> .....	<b>40</b>
5.1 DOMÍNIO E ESCOPO DA ONTOLOGIA DE REFERÊNCIA .....	40
5.2 ENUMERAÇÃO DE TERMOS E CLASSES .....	41
5.3 DEFINIÇÃO DAS PROPRIEDADES E RESTRIÇÕES .....	45
5.4 CRIAÇÃO DAS INSTÂNCIAS.....	47
5.5 CONSIDERAÇÕES DO CAPÍTULO .....	49
<b>6 RESULTADOS</b> .....	<b>50</b>
6.1 MENSAGENS COM CORRELAÇÃO COM ATIVIDADE.....	50
6.2 MENSAGENS COM MARCADORES LINGUÍSTICOS.....	52
6.3 BREVE ANÁLISE DO MAPEAMENTO PROPOSTO.....	53
6.4 CONSIDERAÇÕES DO CAPÍTULO .....	54
<b>7 CONCLUSÃO</b> .....	<b>55</b>

<b>7.1 TRABALHOS FUTUROS .....</b>	<b>55</b>
<b>REFERÊNCIAS.....</b>	<b>57</b>
<b>ANEXO A - Mensagens retiradas de ATV_01_07 – AT3 .....</b>	<b>62</b>
<b>ANEXO B - Termos-Chave relacionados ao domínio da Estatística .....</b>	<b>66</b>
<b>ANEXO C - Mensagens retiradas de ATV_01_07 – AT2 .....</b>	<b>70</b>
<b>APÊNCIDE A - Mensagens significativas retiradas de ATV_01_07 – AT3 .....</b>	<b>74</b>
<b>APÊNDICE B - Mensagens significativas retiradas de ATV_01_07 – AT2.....</b>	<b>77</b>

## 1 INTRODUÇÃO

As organizações buscam inovações para seu ambiente de trabalho, pois tarefas podem ser executadas de forma colaborativa. A colaboração permite que os indivíduos possam compartilhar objetos, informações e ideias na tomada de decisão. No entanto, as informações em um ambiente colaborativo, em sua maioria, são modeladas em diferentes pontos de vista sobre o conhecimento do contexto, formando uma semântica local para cada um deles o que dificulta o processamento de dados (GU *et al.*, 2004).

Em um cenário colaborativo é importante à presença de instrumentos informatizados que possibilitem mensurar e avaliar a colaboração de cada indivíduo, de modo que seja possível identificar se houve colaboração por parte dos participantes e se as informações compartilhadas durante a realização de uma atividade foram realmente significativas (ISHIKAWA *et al.*, 2017).

As ontologias representam e capturam um conhecimento bem fundamentado sobre o domínio. Essa fundamentação possui o intuito de proporcionar um vocabulário comum e bem estruturado desse conhecimento. Mapeamentos ontológicos podem ser utilizados para integração de dados, consulta de respostas e tarefas de análise comparativa entre ontologias (ABBAS; BERIO, 2013).

Existem trabalhos na literatura que se beneficiam do uso de mapeamentos ontológicos como ferramenta de análise e avaliação da colaboração em ambientes virtuais. Um alinhamento entre duas ontologias de origem foi utilizado por Lmat *et al.* (2016) em um estudo de caso para orientação em matemática educacional e concluíram que a técnica com base na estruturação obteve o melhor resultado, porém, o alinhamento feito manualmente ainda é referência quando se trata de minimização de ambiguidades existentes nas ontologias.

O uso de múltiplas ontologias também é abordado por Wimalasuriya e Dou (2009) para a extração de informações. A técnica utilizada pelos autores foi baseada no uso de expressões regulares para capturar informações por meio de palavras específicas, frases e características linguísticas em simples ontologias e múltiplas ontologias. Os resultados trouxeram base para afirmar as vantagens da hipótese do uso de múltiplas ontologias em extração de informações.

Dessa forma, este trabalho realiza o uso de ontologias para um mapeamento usando a metodologia *Ontology Development 101* segundo Noy e McGuinness (2001),

capaz de gerar dados a partir de ontologias de referência mensurando quais mensagens são significativas. Os casos de testes foram obtidos por meio do *Collabora* que é um ambiente que permite a resolução de atividades de forma colaborativa por meio de um *chat* (ISHIKAWA *et al.*, 2017). As mensagens neste ambiente são avaliadas de forma manual pelo professor e isto ocasiona maior tempo de análise.

## 1.1 OBJETIVOS

Criar um mapeamento ontológico de atividades e de interações em um *chat* capaz de mensurar que mensagens são significativas a fim de facilitar a avaliação de colaboração entre os participantes durante a resolução de atividades colaborativas.

Os objetivos específicos deste trabalho são:

- Identificar técnicas e ferramentas para mapeamento ontológico.
- Analisar o funcionamento do ambiente *Collabora* para identificar como as atividades colaborativas são realizadas.
- Realizar testes da ontologia proposta.

## 1.2 ORGANIZAÇÃO DO TRABALHO

Este trabalho é constituído de sete capítulos. O capítulo 2 aborda uma visão geral sobre ontologias, contemplando sua definição em níveis de classificação, mapeamentos ontológicos, algumas ferramentas e metodologias utilizadas para o desenvolvimento de ontologias.

O capítulo 3 relata o estado da arte de Ontologias e Engenharia de software e Ontologias e Mapeamentos. O capítulo 4 apresenta conceitos sobre atividades colaborativas e o ambiente de aprendizagem virtual denominado *Collabora* usando para a realização dos casos de testes.

O capítulo 5 apresenta o mapeamento ontológico proposto para avaliar as mensagens significativas entre os participantes. O capítulo 6 descreve os testes realizados com a ontologia aplicada usando as mensagens obtidas do *Collabora*.

O capítulo 7 apresenta as considerações finais desse trabalho, assim como também identifica possíveis trabalhos futuros.

## 2 ONTOLOGIAS: UMA VISÃO GERAL

Este capítulo descreve uma visão geral sobre ontologias. A seção 2.1 aborda os conceitos utilizados na literatura sobre o tema apresentando os níveis ontológicos, sua classificação, domínio e classes de aplicação. A seção 2.3 relata os conceitos sobre mapeamentos ontológicos. A seção 2.4 aborda as principais ferramentas no tema ontologia. A seção 2.5 descreve algumas metodologias para desenvolvimento de ontologias. Por fim, a última seção apresenta as considerações finais do capítulo.

### 2.1 CONCEITOS

O termo ontologia é utilizado com diferentes significados em diversas áreas. Segundo Guarino e Giaretta (1998), quando tratada na área da filosofia, ontologias são descritas como o ramo que aborda a natureza e estrutura da “realidade”. Aristóteles define ontologia como ciência “ser do ser”, denotada como o estudo dos atributos pertencentes às coisas por conta de sua própria natureza.

Na ciência da computação, Gruber (1993) sugere que ontologia se define por uma especificação explícita de uma conceituação. Além disso, Guarino e Giaretta (1998) complementam essa definição de forma: “uma ontologia é uma especificação explícita e parcial de uma conceituação”. Em síntese, o termo conceituação ainda é definido como uma forma de representação do que é conhecido como universo de discurso.

Em relação a representação, a ontologia é conectada a subárea da inteligência artificial denotada como Processamento de Linguagem Natural (PLN). Esta área estuda como permitir aos seres humanos a comunicação com computadores de maneira eficiente de fácil interação (NORVIG; RUSSEL, 2004).

Segundo Coppin (2004), está se tornando cada vez mais importante que computadores sejam capazes de entender linguagens naturais, tendo em vista a quantidade de dados textuais não estruturados que existem no mundo e em especial na internet, que já alcançou proporções não gerenciáveis. É impraticável para seres humanos realizar buscas em bases de dados apenas com a utilização de técnicas tradicionais, como consultar *booleanas* ou de banco de dados.

Com o desafio de permitir a comunicação entre humanos e máquinas, esforços atuais na área congregam objetivos que, segundo Liddy (2003), são usuais no PLN (Processamento de Linguagem Natural). Tais objetivos se dão pela recuperação de informação por meio de textos, tradução automática, interpretação de textos e realização de inferências. A construção de ontologias a partir de textos, por exemplo, auxilia no alcance dos objetivos do PLN. Algumas técnicas e ferramentas podem ser utilizadas para análise e interpretação de uma linguagem natural, porém, é importante frisar a distinção entre tais técnicas e ferramentas quando aplicadas na linguagem falada ou na linguagem escrita.

De acordo com Coppin (2004), quando abordado o ponto de vista linguístico, a PLN pode ser dividida em cinco níveis:

- **Fonologia:** Relacionada ao reconhecimento dos sons que compõe as palavras de uma língua.
- **Morfologia:** Capacidade de reconhecer palavras em termos das unidades primitivas que a compõem e examinam os modos pelos quais palavras podem se desmembrar em componentes e como isto afeta o seu *status* gramatical.
- **Sintaxe:** Determina o papel de cada palavra em uma sentença, permitindo que um sistema computacional converta sentenças em uma estrutura que possa ser mais facilmente manipulada. Neste estágio, são aplicadas regras gramaticais na linguagem que estiver sendo utilizada.
- **Semântica:** Consiste em examinar e associar um significado a uma estrutura sintática, pois é possível a sentença estar sintaticamente correta e semanticamente sem sentido. É desejável que um sistema computacional seja capaz de entender sentenças com sintaxe incorreta, mas que transmitam informação útil de um ponto de vista semântico. Nível em que as ontologias se encontram para representar conhecimento.
- **Pragmática:** Responsável por verificar se o significado que foi associado a uma estrutura é de fato o significado mais apropriado em um considerado contexto.



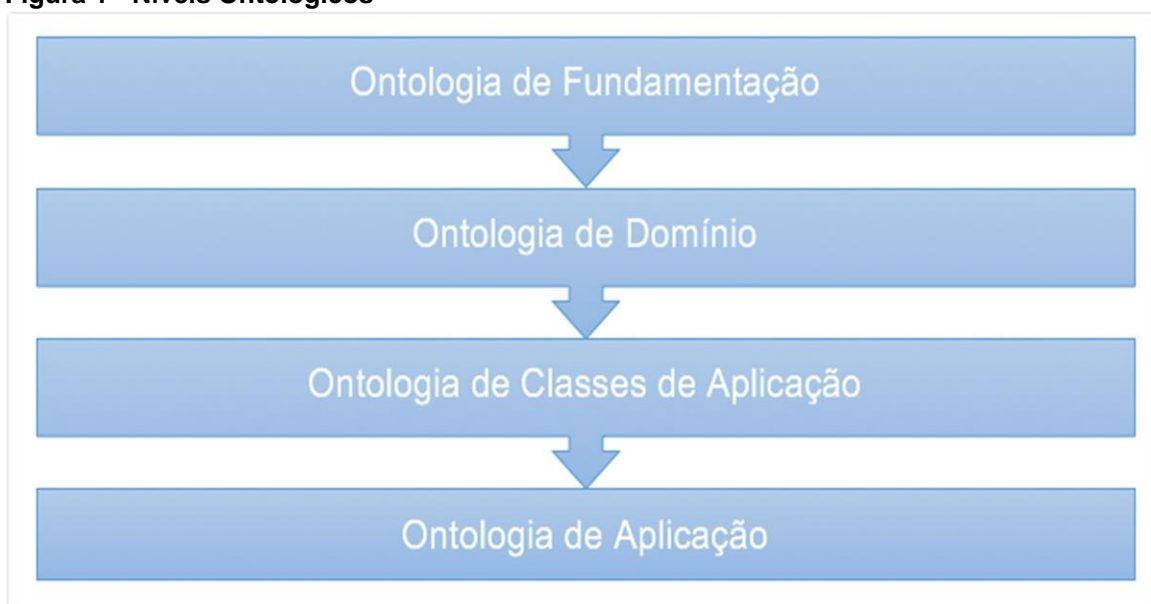
Para sintetizar, além destes níveis de análise, sistemas de Processamento de Linguagem Natural necessitam aplicar algum tipo de conhecimento sobre o mundo. Neste contexto, as ontologias como definidas no início desta seção possuem um papel importante no processo de representação de conhecimento para auxílio do processamento de linguagens naturais.

## 2.2 NÍVEIS ONTOLÓGICOS

Existe uma classificação básica entre dois tipos de Ontologias. As ontologias de alto nível ou de fundamentação e as de baixo nível. Segundo Degen (2001), as ontologias de fundamentação são teorias ou especificações de categorias de entidades gerais, independentes de domínio, como eventos, relações, tempo, espaço, entre outros. Enquanto que as ontologias de baixo nível são de acordo com Guarino (1998) atreladas ao domínio particular do que se deseja representar. Este tipo possui entidades específicas do domínio e podem ser divididas em ontologias de: domínio, classe de aplicações e de aplicações.

Uma visão geral da classificação de níveis e como esses níveis são organizados está ilustrado na Figura 1.

**Figura 1 - Níveis Ontológicos**



**Fonte: Adaptado de Guarino (1998)**

As Ontologias de Fundamentação são livres de domínio, abrangem teorias de partes, teorias de todo-parte, diferença entre tipos e instâncias, rigidez e anti-rigidez, entre outros. Dentre as ontologias de fundamentação mais conhecidas, estão as descritas no Quadro 1.

**Quadro 1 - Ontologias de Fundamentação**

<b>Ontologia</b>	<b>Autor(es)</b>	<b>Ano</b>
SUMO ( <i>Suggested Upper Merge Ontology</i> )	NILES; PEASE	2001
UFO ( <i>Unified Foundational Ontology</i> )	GUIZZARDI	2005
GFO ( <i>General Formal Ontology</i> )	HERRE <i>et al.</i>	2006
DOLCE ( <i>Descriptive Ontology for Linguistic and Cognitive Engineering</i> )	BOTTAZI; FERRARIO.	2008

**Fonte: Autoria própria**

Guizzard e Wagner (2005) criaram uma proposta de fundamentação unificada que fornece uma fundamentação ontológica para a construção de modelos conceituais. A chamada UFO (do inglês, *Unified Foundational Ontology*), é dividida em três categorias denominadas: UFO-A, UFO-B e UFO-C.

UFO-A (*Ontology of Endurants*), define o núcleo da ontologia UFO por meio de termos relacionados a aspectos estruturais como conceitos gerais de objetos, suas propriedades intrínsecas e relacionais, os tipos que instanciam, papéis que podem representar e etc. Em outras palavras, esta categoria trata conceitos do tipo duradouros (*Endurants*) independentes de domínio.

UFO-B (*Ontology of Perdurants*), como complemento da UFO-A, este tipo de ontologia define termos relacionados a processos, eventos e suas relações.

UFO-C (*Ontology of Social and Intentional Entities*), é definida como um complemento a UFO-B, tratando de termos relacionados à intencionalidade e as entidades sociais.

Uma ontologia de domínio é capaz de capturar o conhecimento válido para um tipo particular de domínio. Por exemplo, domínios como medicina, mecânica, biologia, estatística, entre outros. Este tipo de ontologia expressa conceituações de domínios particulares descrevendo o vocabulário relacionado a um domínio genérico, tal como a Estatística. Guarino (1998) relata que “uma ontologia de domínio apresenta um vocabulário relativo a um vocabulário genérico”.

Segundo Guizzardi (2005), dentre os objetivos de uma ontologia de domínio estão:

- Compartilhar informação;
- Reusar elementos do domínio;
- Tornar suposições do domínio explícitas;
- Separar conhecimentos de domínio de conhecimento operacional e analisar o conhecimento de domínio;
- Promover interoperabilidade entre aplicações;
- Produzir modelos conceituais simples e com consistência lógica.

As ontologias do tipo de Classe de Aplicações descrevem conceitos que se fundam de domínios e tarefas específicas. Estes tipos de ontologia são especializações de ontologias de domínio e tarefa. De acordo com Guizzard (2005), esses conceitos frequentemente correspondem a entidades do domínio e aos papéis executados por eles enquanto estão executando uma tarefa.

As especializações de conhecimento são criadas com o intuito de incorporarem conhecimento que são dependentes de tarefas e suas aplicações. Possuem um propósito exclusivo e sua funcionalidade necessita ser executada por algum sistema específico ou por uma classe de sistemas. A ideia de ontologias de aplicação é semelhante ao processo de especialização de um conhecimento em um determinado domínio (FALBO *et al.*, 2002).

### 2.3 MAPEAMENTO ONTOLÓGICO

Um mapeamento ontológico pode ser visto como um provedor de soluções no cenário de pesquisas baseadas em ontologias (CHOI *et al.*, 2006). As ontologias são vistas como um fator chave para a solução da interoperabilidade entre sistemas heterogêneos e aplicações *web* semântica. Com isto, o mapeamento ontológico se torna necessário para realizar a combinação de simples ou múltiplas ontologias.

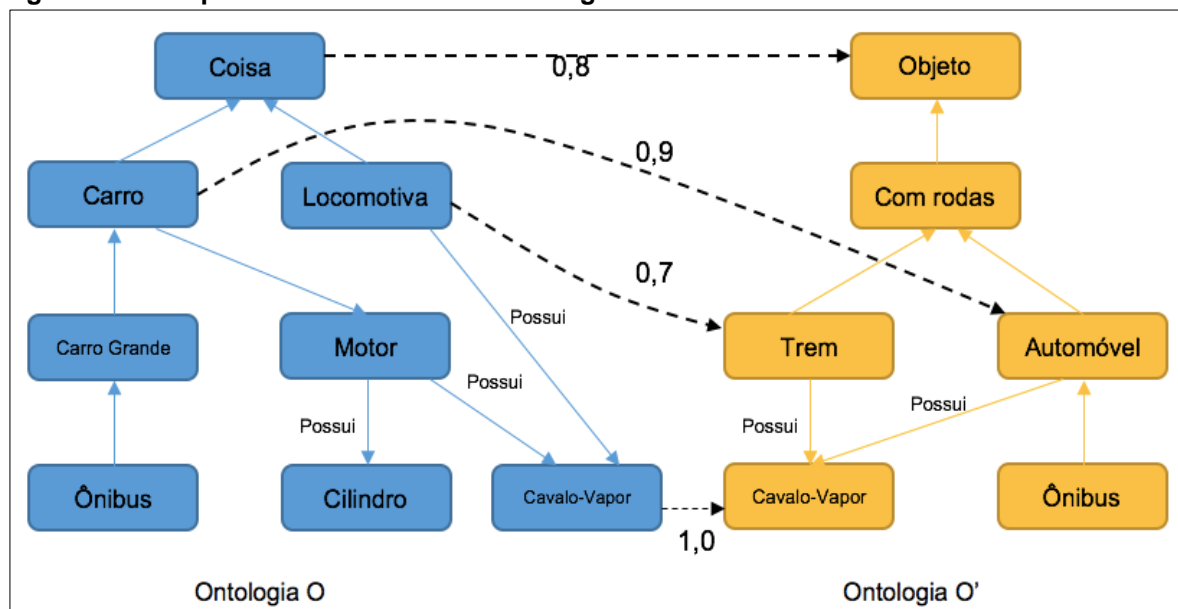
Mapeamentos ontológicos são comumente divididos em três categorias: mapeamento entre uma ontologia integrada global e locais, mapeamento entre ontologias locais e fusão e alinhamento de ontologias (CHOI *et al.*, 2006). A primeira categoria de mapeamento suporta a integração de uma ontologia descrevendo a relação entre uma ontologia global integrada e ontologias locais enquanto que a

segunda permite a interoperabilidade para ambientes dinâmicos e distribuídos como mediador entre dados distribuídos. Por fim, a terceira categoria pode ser utilizada como alinhamento entre ontologias para processo de reuso das mesmas.

De modo geral, duas ou mais ontologias podem estar relacionadas e essas relações podem ser nomeadas por meio de funções. Segundo Kalfoglou e Shorlemmer (2003), um alinhamento de ontologias é a tarefa de estabelecer uma coleção de relações binárias entre vocabulários de duas ontologias. Essa relação binária pode ser composta por um par de funções totais a partir de uma fonte intermediária comum.

Um exemplo é o uso de palavras-chave como função de comparação, comumente o alinhamento é feito por meio de duas ontologias  $O_1$  e  $O_2$ , sendo uma ontologia de origem e outra para alinhamento. O alinhamento é realizado a partir de combinações dadas a partir de uma função inserida e com isto é possível gerar uma ontologia resultante, ou seja, um mapeamento ontológico de duas ontologias. Um exemplo está ilustrado na Figura 2.

Figura 2 - Exemplo de Alinhamento de Ontologias



Fonte: Adaptado de Abolhassani *et al* (2006)

Neste exemplo, pode-se observar duas ontologias: a Ontologia O e a O', onde as arestas pontilhadas indicam combinações entre os elementos das ontologias. Além disso, é possível indicar a confiança ou força desse alinhamento por meio da

atribuição de peso às arestas. Verifica-se essa atribuição na correspondência entre Carro e Automóvel com 0.9.

## 2.4 FERRAMENTAS PARA MAPEAMENTO

Algumas ferramentas para mapeamento ontológico foram analisadas por Choi (2006), com a utilização de outros critérios como: estratégia de mapeamento ou algoritmo, conhecimento estruturado, léxico e de domínio. Um resumo dos resultados obtidos por esta comparação é apresentado no Quadro 2. Ressalta-se que este comparativo apresenta apenas ferramentas que utilizam duas ou mais ontologias para mapeamento.

**Quadro 2 – Comparativo entre ferramentas de mapeamento de ontologias segundo Choi**

<b>Ferramenta</b>	<b>Glue</b>	<b>MaFra</b>	<b>Lom</b>	<b>Prompt</b>	<b>Fca-merge</b>
<b>Entrada</b>	Duas taxonomias com seus dados instanciados em ontologias	Duas Ontologias	Duas listas de termos de duas ontologias	Duas entradas ontológicas	Duas entradas ontológicas e um conjunto de conceitos das ontologias em um documento
<b>Saída</b>	Um conjunto de pares de conceitos similares	Mapeamento de duas ontologias através da semântica de ponte	Lista de pares combinados de termos com pontuação e ranking de semelhança	Uma ontologia mesclada	Uma ontologia mesclada
<b>Estratégia de Mapeamento ou Algoritmo</b>	Aprendizagem de máquina	Ponte semântica	Semelhança lexical	Combinações linguísticas, inferências baseada em heurísticas	Heurística baseada em análises

Fonte: Adaptado de Choi (2006)

Ainda segundo Alatrish (2012), algumas das ferramentas mais adotadas para a criação e manipulação de ontologias são: *Apollo*, *OntoStudio*, *Protégé*, *Swoop* e *TopBraid Composer*. Este autor realizou uma comparação entre as ferramentas usando os critérios como: generalização, expressividade, complexidade, documentação, escalabilidade, importação e exportação, entre outros. Uma adaptação de comparações do trabalho é apresentada no Quadro 3 contendo apenas critérios que, são de alguma forma, relevantes para o embasamento deste trabalho.

Quadro 3 – Comparativo entre ferramentas de ontologias segundo Alatrish

Ferramenta	Apollo	OntoStudio	Protégé	Swoop	TopBraid Composer
Desenvolvedor	KMI (Open University)	OntoPrise	SMI (Stanford University)	MND (University of Maryland)	TopQuadrant
Licença	Open Source	Licença de Software	Open Source	Open Source	Licença de Software
Importação/Exportação	Sim	Sim	Sim	Sim	Sim
Armazenamento de Ontologia	Files	DBSM	File and DBMS (JDBC)	HTML	DBMS
Gerenciamento de Backup	Não	Não	Não	Não	Sim
Trabalho Colaborativo	Não	Sim	Sim	Sim	Sim
Biblioteca de Ontologias	Sim	Sim	Sim	Não	Sim

Fonte: Adaptado de Alatrish (2012)

As ferramentas *Apollo*, *Protégé* e *Swoop* possuem licença *open source* e foram desenvolvidas por estudos em universidades, enquanto que a *OntoStudio* e *TopBraid Composer* necessitam de licença para o uso. Além disso, algumas aplicações das ferramentas necessitam de conhecimento de linguagens em particular, por exemplo, o *Swoop*.

Todas as ferramentas possuem importação e exportação e algum tipo de armazenamento de ontologias. Apenas a ferramenta *Swoop* não possui uma biblioteca de ontologias o que deixa o uso da aplicação um pouco mais difícil. A seguir explica-se a ferramenta *Protégé*, pois será a utilizada no desenvolvimento deste trabalho.

#### 2.4.1 Protégé

O *Protégé*, além de um editor, trata-se de um framework para obtenção de conhecimento e possui licença *open source*. A ferramenta, ainda possui suporte para extensão de funcionalidades por meio de *plug-ins* e também provê uma base flexível para o desenvolvimento de protótipos e aplicações de maneira eficiente (NOY;

MCGUINNESS, 2006). A arquitetura do *Protégé* é modulada e permite a inserção de novos recursos. Atualmente importa/exporta para diversas linguagens de ontologias, dentre elas a OWL, OIL, XML, RDF, Flogic e Prolog são alguns exemplos.

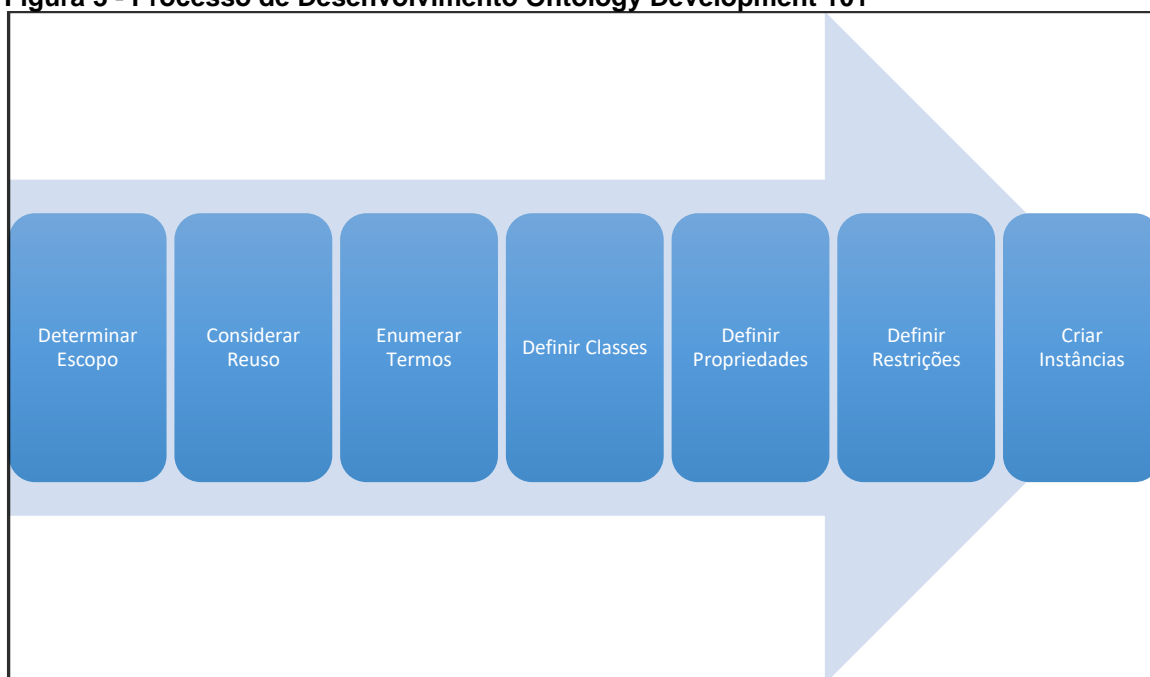
A ferramenta foi desenvolvida por Mark Musen e seu grupo de Informática Médica da Universidade de Stanford em 1999. De acordo com Gasevic *et al* (2009), o *Protégé* é a principal ferramenta de engenharia ontológica na literatura. De acordo com o site oficial do *Protégé*, a ferramenta conta com mais de 300.000 usuários registrados.

## 2.5 ESTRATÉGIAS DE MAPEAMENTO

Esta seção descreve conceitos gerais e específicos sobre tipos metodologias e estratégias de mapeamento com suas principais características. As subseções apresentam os processos utilizados no desenvolvimento de ontologias em cada metodologia proposta.

### 2.5.1 Ontology Development 101

Segundo Noy e McGuinness (2001), não existe um método correto ou metodologia de desenvolvimento para ontologias. Por isso, os autores da Universidade de Stanford criadora da ferramenta *Protégé* sugerem um processo denominado *Ontology Development 101*. Este processo consiste em sete passos ilustrados na Figura 3.

**Figura 3 - Processo de Desenvolvimento Ontology Development 101**

**Fonte: Adaptado de Noy e McGuinness (2001)**

O primeiro passo, Determinar Escopo, é usado para determinar o domínio e o escopo de uma ontologia, identificando de forma clara o propósito e os cenários de utilização da ontologia a ser desenvolvida. Para isto, é preciso responder alguns exemplos de questões, segundo Noy e McGuinness (2001):

- O que abrange o domínio da ontologia?
- Para que se utilizará a ontologia?
- Quais questões a ontologia deverá responder?
- Quem utilizará e manterá a ontologia?

O próximo passo, Considerar Reuso, é importante porque antes do desenvolvimento de uma ontologia é aconselhável verificar a existência de ontologias no domínio a ser trabalhado e que podem ser reutilizadas em um projeto.

Enumerar termos, terceiro passo, permite relacionar uma lista de termos importantes presentes no ambiente de domínio da ontologia a ser criada. A relação de termos torna-se importante para os passos subsequentes.

Através da lista de termos do passo anterior é possível definir as classes de domínio e hierarquia das mesmas. Este passo é responsável por extrair os objetos que representam uma ou mais classes. A partir de um conjunto de classes definido, é



necessário organizar as classes de forma hierárquica por meio de um nível de abstração mais generalizado em direção as classes mais específicas.

O quinto passo, Definir Propriedades, utiliza a lista dos termos remanescentes e deve-se observar se eles correspondem as propriedades de dados ou de relações de classe para uma determinada classe.

Definir Restrições, sexto passo, permite verificar se uma propriedade de classe é no formato de dados, tais como *string* ou número, por exemplo. Caso contrário, se a propriedade for do tipo relação deve-se definir a que classe a relação se refere. Considera-se restrições sobre cardinalidade e valores válidos para as propriedades.

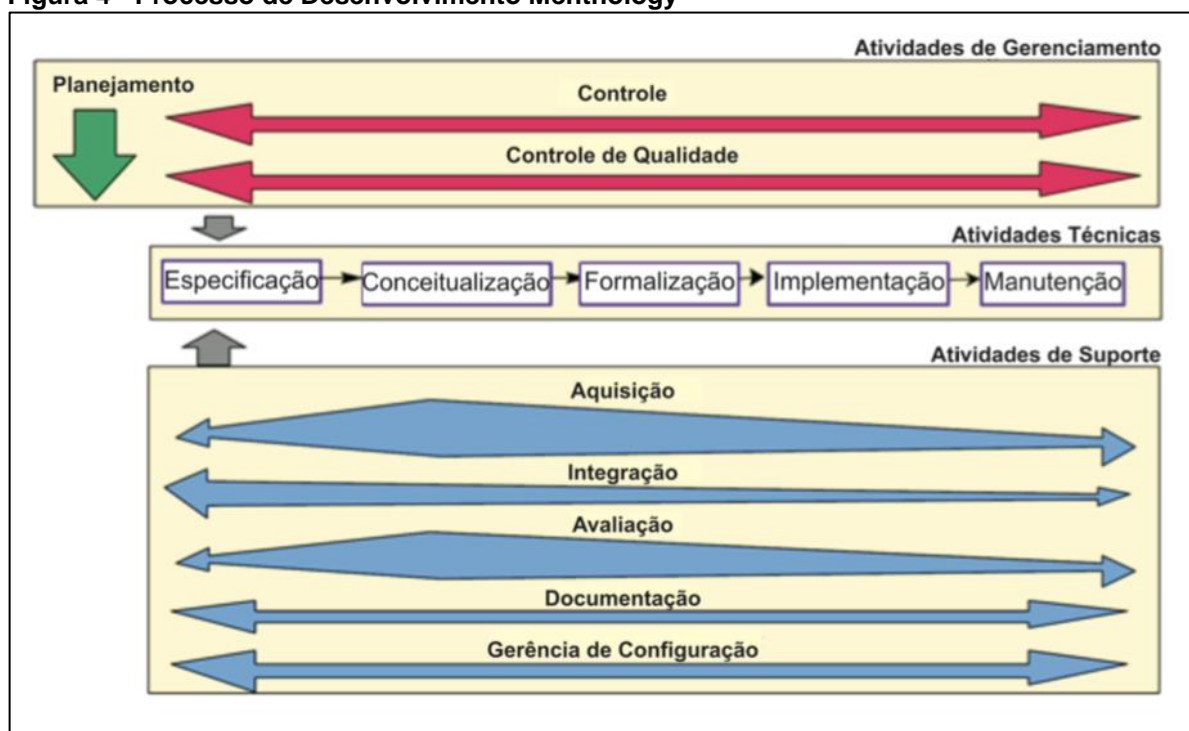
Por fim, o último passo, Criar Instâncias, permite a criação de instâncias da ontologia a partir da definição das classes, atribuindo valores as suas propriedades de dados e relações.

A metodologia *Ontology Development 101* trata de um guia de quais passos devem ser seguidos para o desenvolvimento de uma ontologia. Por se tratar de um guia para a ferramenta *Protégé*, comentada na seção anterior e escolhida para utilização neste trabalho, esta metodologia é a que melhor se enquadra na aplicação da proposta e será detalhada no capítulo 6.

### 2.5.2 Mentology

A *Mentology* é uma metodologia de desenvolvimento de ontologias baseada na construção de ontologias a partir do conhecimento de um domínio. Foi idealizada por um grupo de pesquisa em Engenharia de Ontologias da Universidade Politécnica de Madri. Suas principais atividades são: especificação de requisitos, conceitualização do domínio de conhecimento, formalização do modelo conceitual de uma linguagem formal, implementação de um modelo formal e manutenção de ontologias já implementadas. O processo de desenvolvimento desta metodologia está ilustrado na Figura 4.

Figura 4 - Processo de Desenvolvimento Mentholology



Fonte: Jones *et al* (1998)

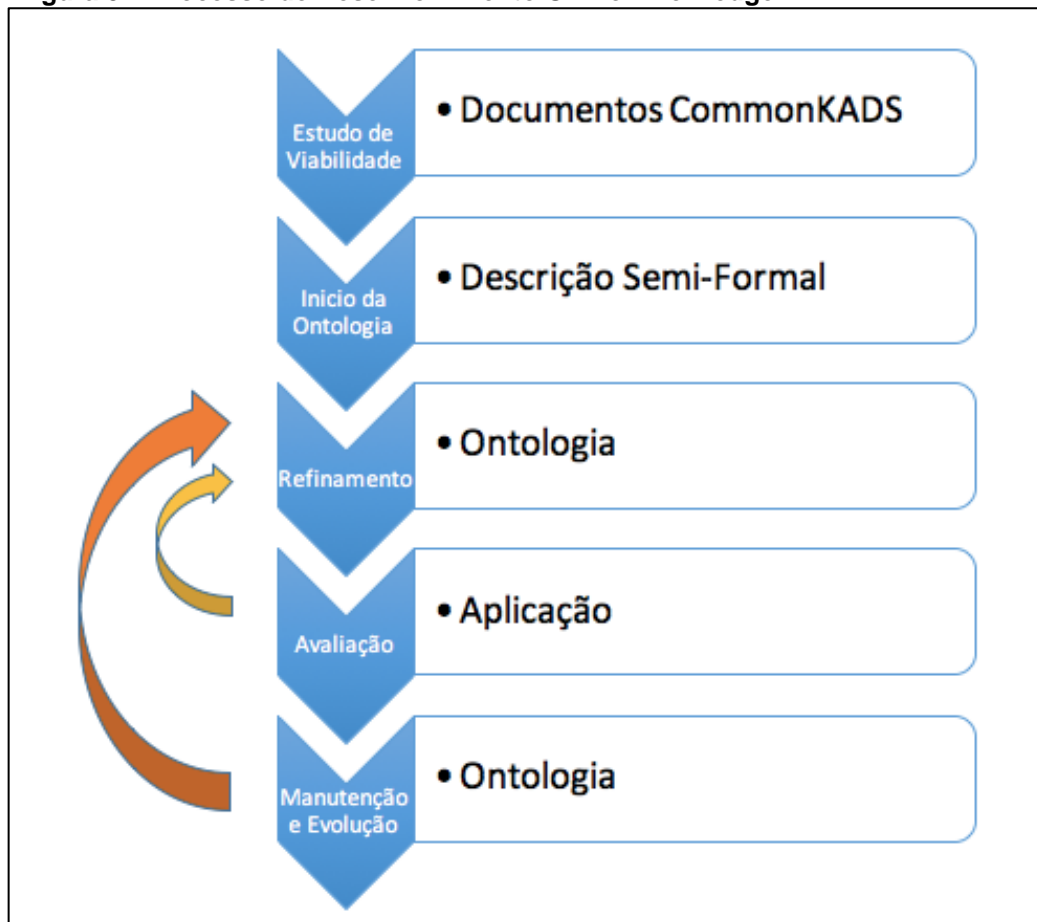
Como pode ser observado na Figura 4, a *Mentholology* provê um ciclo de vida baseado no desenvolvimento de uma ontologia e isto acontece de acordo com o processo de evolução da criação da ontologia. Esta metodologia possui ainda atividades para suporte desempenhadas durante o processo de construção de uma ontologia. Estas atividades são a aquisição de conhecimento, integração, avaliação, documentação e gerenciamento de configuração.

Este processo é inspirado em preceitos do ciclo de vida de um software, as atividades privilegiadas nesta metodologia são destinadas ao gerenciamento, desenvolvimento e suporte ao ciclo de vida de uma ontologia. São realizadas três divisões para a formalização e suporte da ontologia, são elas as: i) atividades de gerenciamento, ii) atividades técnicas e iii) atividades de suporte. Cada etapa possui um ciclo de funcionamento e cada etapa se relaciona com a outra de forma a completar o ciclo total de desenvolvimento de uma ontologia com o método *Mentholology*.

### 2.5.3 On-To-Knowledge

Esta metodologia foi criada com o intuito de desenvolvimento de ontologias no âmbito de Sistemas de Gestão de Conhecimento. O processo de desenvolvimento proposto é dividido em cinco fases, conforme mostra a Figura 5.

**Figura 5 – Processo de Desenvolvimento On-To-Knowledge**



**Fonte: Adaptado de Fensel (2018)**

A fase de estudo de viabilidade também é chamada de *kick-off*, nesta etapa os requisitos para a construção da ontologia são capturados e especificados, questões de competência são identificadas, ontologias reutilizáveis são estudadas e uma primeira versão da ontologia é construída na fase de início da ontologia na qual há a descrição semi-formal da mesma.

No passo de refinamento uma nova ontologia é construída a partir da primeira versão, ou seja, o refinamento de fato.

A fase de avaliação é responsável por validar os requisitos e as questões de competência são checadas para que a ontologia possa ser colocada em um ambiente de desenvolvimento.

Por fim, a manutenção refere-se às atividades de adaptação da ontologia às mudanças que podem ocorrer nos requisitos e auxilia na correção de erros.

## 2.6 CONSIDERAÇÕES DO CAPÍTULO

Este capítulo apresentou os conceitos associados às ontologias e suas principais características segundo autores especialistas da área.

Foi descrito também uma breve introdução sobre os níveis ontológicos existentes e suas classificações e abordou uma visão geral sobre o que são mapeamentos ontológicos. A prática de mapeamentos ontológicos implica diversos campos como o aprendizado de máquinas, teorias formais para heurísticas, redes conceituais, esquema de banco de dados, linguística e entre outros (ALATRISH, 2012).

Constatou-se que as ontologias de domínio são as mais adequadas como estratégia de descrição do conhecimento sobre os elementos essenciais considerados na aprendizagem colaborativa envolvendo pessoas e artefatos do sistema.

Por fim, foram apresentadas algumas metodologias de desenvolvimento para ontologias com suas principais características e etapas, sendo que a metodologia *Ontology Development 101* será utilizada neste trabalho e seus artefatos foram implementados na ferramenta *Protégé*.

### 3 ESTADO DA ARTE

Este capítulo descreve o estado da arte sobre ontologias realizado por meio de pesquisa bibliográfica usando sites de busca como Google Scholar, ACM, IEEE, entre outros. A seção 3.1 abrange a ontologia e desenvolvimento de software. A seção 3.2 ontologia e mapeamentos, bem como relata alguns trabalhos relacionados. Por fim, a última seção apresenta as considerações finais do capítulo.

#### 3.1 ONTOLOGIA E DESENVOLVIMENTO DE SOFTWARE

Como consequência do desafio da manipulação de dados processados, isto é, informação, várias soluções foram propostas na tentativa de automatizar e integrar fontes de informação distribuídas. Entre estas, as tecnologias ligadas à semântica têm atraído mais atenção quando o assunto é melhorar o cenário (SHVAIKO; EUZENAT, 2013).

A superação da heterogeneidade semântica é alcançada a partir de dois passos: entidades correspondentes para determinar um possível alinhamento, ou seja, um conjunto de correspondências e interpretação de um alinhamento de acordo com as necessidades de uma aplicação específica, como tradução de dados, consulta e resposta ou combinações.

Alguns dos desafios para os próximos anos na Engenharia de Software apontado por diversos autores, segundo Isotani *et al* (2015), é a integração do conhecimento em ambientes de desenvolvimento de software seguido da dificuldade de refinamento e gerenciamento eficaz das fases de desenvolvimento de software, tais como: requisitos, planejamento e implementação e à crescente demanda de desenvolvimento de sistemas que manipulam, armazenam, processam e apresentam a informação presente na *Web*.

Quando o assunto é integração do conhecimento, o termo “conhecimento” faz referência aos métodos, processos e padrões necessários para planejar e coordenar o desenvolvimento de um software. Muito tempo é desperdiçado pela falta de formalização explícita do conhecimento e de formas de compartilhamento no ramo da Engenharia de *Software*.

Devido a essa carência, há ainda a dificuldade de refinamento e gerenciamento eficaz das fases de desenvolvimento de um *software*. Quanto maior a complexidade de um sistema, maiores serão as dificuldades e custos para gerenciar as fases de desenvolvimento do mesmo (COSKIN; GRABOWSKI, 2005).

Existe a necessidade de se construir ferramentas e métodos avançados mais inteligentes para suporte às diversas atividades da Engenharia de Software, para que as mesmas possam diminuir o custo e tempo de desenvolvimento de aplicações, além de propiciar maiores benefícios a desenvolvedores podendo garantir mais qualidade em um produto final.

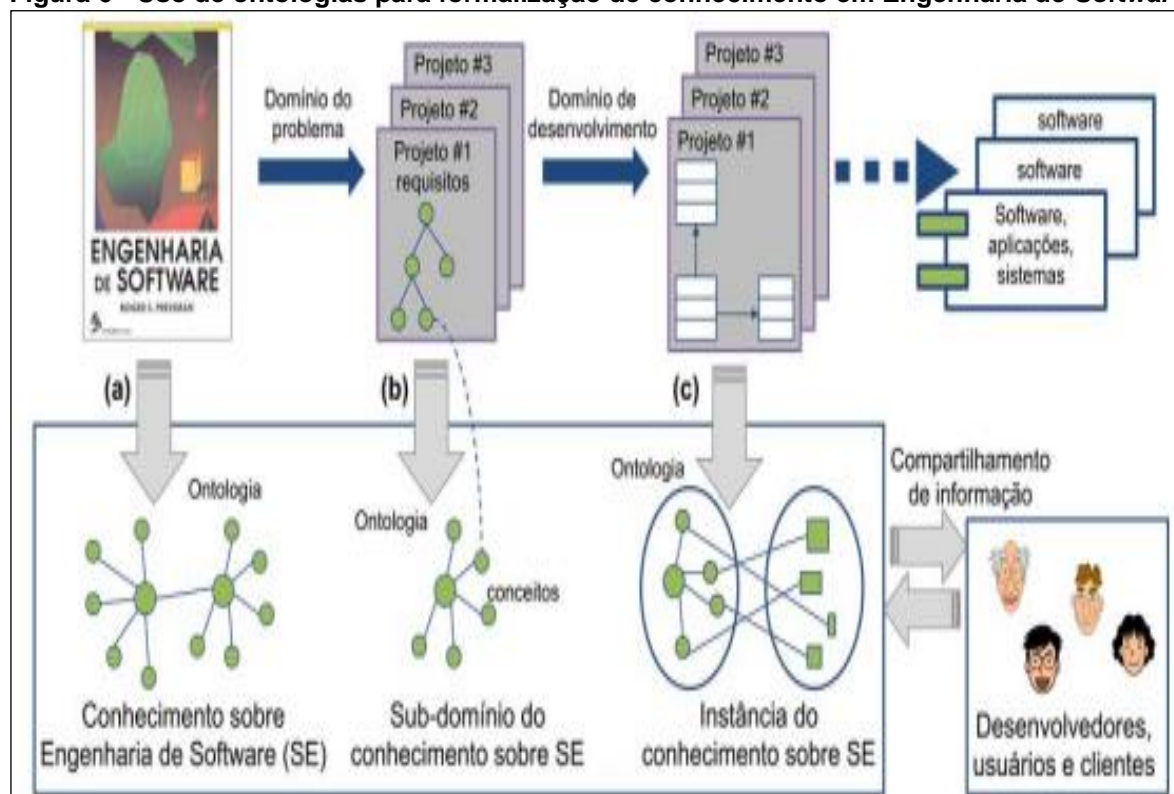
Relacionado à crescente demanda de desenvolvimento de sistemas que manipulam, armazenam, processam e apresentam informações presentes na *Web*, existe a necessidade de criar programas inteligentes e interativos que possibilitem melhorar os processos de gerenciamento e tomada de decisão em aplicações. Neste contexto, é fundamental promover uma abordagem de forma interdisciplinar para gerar novos métodos, processos e ferramentas de desenvolvimento de software.

Entre as propostas para enfrentar os desafios e dificuldades abordadas, o uso de ontologias em Engenharia de *Software*, e outras tecnologias ligadas a *Web Semântica*, segundo Cesare *et al* (2009), possuem grande potencial de impacto nos diversos aspectos de desenvolvimento de software como, por exemplo, na criação de modelos e linguagens mais apropriadas para representar o software. Estes modelos podem auxiliar no processo de análise e especificação de requisitos, na organização do projeto de desenvolvimento de software, na verificação e validação de sistemas. Apesar do assíduo interesse em utilizar ontologias para apoiar diferentes fases do processo de desenvolvimento de software, ainda há uma carência de trabalhos que discutam oportunidades e desafios constantes na área (CESARE *et al*, 2009).

De acordo com Isotani *et al* (2015) sobre o uso de ontologias na Engenharia de Software, tem-se que (Figura 6): (a) ao que se diz sobre “conhecimento” na área de Engenharia de *Software* como, por exemplo, os diversos processos de desenvolvimento existentes, tipos de artefatos e como estes artefatos e fases de processo se relacionam, podem ser representados utilizando uma ontologia de domínio; (b) representa o domínio de um problema, ou seja, os requisitos também podem ser formalizados por meio do uso de ontologias de domínio que são capazes de representar informações específicas dos conceitos relativos a um domínio da Engenharia de Requisitos; e (c) o espaço de uma solução pode fazer o uso de

ontologias de domínio para conter, por exemplo, os conceitos relativos ao projeto arquitetural de um software com influência do compartilhamento de informação por meio de desenvolvedores, usuários e clientes envolvidos.

**Figura 6 - Uso de ontologias para formalização de conhecimento em Engenharia de Software**



Fonte: Isotani *et al* (2015)

### 3.2 ONTOLOGIA E MAPEAMENTO

O progresso das tecnologias da informação e da comunicação propiciou uma disparidade de informações em bases de dados. Há um problema na gestão da heterogeneidade entre vários recursos de crescentes informações. Por exemplo, a maioria das pesquisas em um banco de dados com auto avaliação reportam relatórios que já reconhecem o problema da questão de heterogeneidade semântica, isto é, a dificuldade de manipular variações de significados ou ambiguidades na interpretação de entidades continua em aberto como um desafio na tecnologia da informação.

Os interessados em se apoiar no uso de mapeamentos ontológicos para a resolução de problemas são muitas vezes confrontados com a extensa quantidade de trabalhos provenientes de comunidades que reivindicam algum tipo de relevância

quando o assunto é mapeamento de ontologias (KALFOGLOU; SHORLEMMER, 2003). Parte deste problema se dá pela falta de pesquisas abrangentes, terminologias padrões, escassez de métricas de avaliação, entre outros. Por exemplo, na literatura é possível encontrar trabalhos que afirmam a relevância de técnicas que incluem alinhamento, fusão, articulação, integração, morfismo, entre outros. Devido a esta diversidade, torna-se difícil a identificação de quais técnicas ou ferramentas utilizar em específicos projetos ou aplicações.

O mapeamento ontológico é a solução para o problema das heterogeneidades semânticas. Esta área emergente já possui avaliações recentes que indicam que o campo de mapeamento ontológico fez uma melhora mensurável no que se diz respeito ao auxílio de processamento de informações (SHVAIKO; EUZENAT, 2013).

### 3.2.1 Trabalhos Relacionados

Dentre alguns trabalhos relacionados ao mapeamento ontológicos encontrados na literatura. Em 2001, Leidig apresentou uma ontologia que possuía um modelo didático para a caracterização de grafos conceituais e servidores como uma espécie de vocabulário e gramática para informação semântica sobre objetos de conhecimento. O autor considerou a ementa e os conteúdos de um curso como objetos de conhecimento para a sua proposta. O modelo proposto por Leidig tornava explícito o conhecimento didático de forma independente do domínio de conhecimento. O modelo consistia em conceitos didáticos, conjunto de relações e padrões que descreviam usos típicos de conceitos e suas relações.

Uma ontologia CSCL baseada na Teoria da Atividade foi apresentada por Barros *et al* (2002). A proposta foi criada a partir da pesquisa sobre mecanismos de representação para relacionar e integrar elementos de aprendizagem colaborativa existentes em plataformas de aprendizagem. A ontologia CSCL apresentada considerava e relacionava os elementos de aprendizagem, utilizam os mesmos para a definição de novos cenários de aprendizagem colaborativas. Pela proposta de Barros ser baseada na Teoria da Atividade, seus principais componentes são: ferramentas, regras, divisão de trabalho, comunidade, assunto, objeto e resultado.

Como já citado anteriormente neste trabalho, Wimalasuriya e Dou (2009) utilizaram o uso de múltiplas ontologias para extração de informações com base na



utilização de expressões regulares para cruzar combinações com palavras-chave, frases e características em ontologias simples e múltiplas no domínio da extração de informações. Um comparativo entre as vantagens e desvantagens do uso de ontologias múltiplas e ontologias simples foi realizado e pôde ser visto que o sistema com uso de múltiplas ontologias mostrou melhorias em todas as três medidas utilizadas no trabalho para correspondência de subdomínio.

Lmat *et al*, em 2016 utilizou técnicas de estruturação para o alinhamento entre duas ontologias de origem no domínio da matemática educacional. O trabalho obteve resultados suficientes para concluir que o alinhamento de ontologias feito de forma manual ainda é referência em comparação a ferramentas existentes na literatura para este processo, como a ferramenta *SemMatcher*, por exemplo.

Rocha (2017) propôs em sua tese de doutorado uma proposta de avaliação de aprendizagem utilizando uma abordagem qualitativa baseada em mapas conceituais, ontologias e algoritmos genéricos. A proposta utilizou-se de um protótipo de ambiente para ensino, aprendizagem e avaliação da aprendizagem, fundamentado na aprendizagem significativa com a inclusão de editores de mapas conceituais, ontologias e um módulo avaliador. Além disto, o autor utiliza algoritmos genéticos para a avaliação e acompanhamento qualitativo da aprendizagem.

### 3.3 CONSIDERAÇÕES DO CAPÍTULO

Este capítulo abordou o estado da arte ao que se refere a ontologias na engenharia de software e mapeamento por meio de ontologias. Foi apresentada uma breve pesquisa dos principais desafios e usos da ontologia na área de desenvolvimento de software.

Alguns trabalhos relacionados com o uso de ontologias simples e múltiplas para aplicação em ambientes colaborativos ou apenas com o intuito de extração de informação foram apresentados.

O próximo capítulo aborda o tema de atividades colaborativas com suas principais características, ferramentas, conceitos e o estudo de caso deste trabalho, o *Collabora*.

## 4 ATIVIDADES COLABORATIVAS

Este capítulo aborda os conceitos gerais sobre atividades colaborativas, suas principais ferramentas e o ambiente *Collabora*. A Seção 4.1 abrange os conceitos gerais e importantes da colaboração, tal como o modelo 3C: Comunicação, Cooperação e Coordenação. A seção 4.2 apresenta ferramentas utilizadas para colaboração como *chats*, fóruns e ambientes virtuais de aprendizagem. A seção 4.3 relata o ambiente de aprendizagem virtual denominado *Collabora* com suas principais características e por meio do qual as questões foram utilizadas para realização dos casos de teste. Por fim, a última seção descreve as considerações finais do capítulo.

### 4.1 CONCEITOS GERAIS

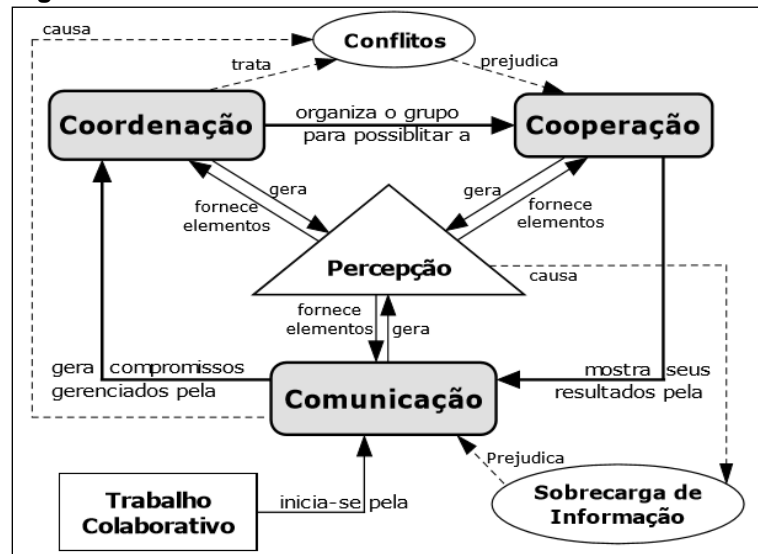
O trabalho em grupo, ou seja, a atividade realizada de forma colaborativa pode auxiliar no aumento da produtividade de indivíduos. Ambientes Colaborativos é a tradução adotada no Brasil para *Groupware* e CSCW (do inglês, *Computer Supported Cooperative Work*), e é caracterizado pela interação entre duas ou mais pessoas com o intuito de realizar um objetivo comum específico.

Os Ambientes Colaborativos de Aprendizagem visam proporcionar um espaço de construção coletiva de conhecimento, em que cada participante se torna autor e colaborador dos resultados gerados pelo grupo na resolução de uma atividade (HAGUENAUER *et al.*, 2007).

Um conceito importante relacionado as atividades colaborativa é a Aprendizagem Colaborativa com Suporte Computacional (CSCL). Trata-se de uma área emergente das ciências de aprendizagem que estuda como pessoas podem aprender em um grupo com o auxílio de computadores (STAHL *et al.*, 2006).

Segundo Fuks (2001), para que ocorra a colaboração o diagrama dos 3C (Comunicação, Coordenação e Cooperação) consegue representar de forma intuitiva o que é necessário e como ocorrem as fases de um processo de colaboração. O diagrama da Figura 7 apresenta a estruturação do modelo 3C difundido na literatura.

Figura 7 - Estrutura do Modelo 3C



Fonte: Fuks *et al* (2001)

Na etapa de Comunicação os indivíduos buscam construir entendimento comum por meio de consenso compartilhando ideias, discutindo e negociando tomadas de decisões. Para Fussel *et al.* (1998), nesta etapa os indivíduos devem se comunicar com o objetivo de conseguir desenvolver tarefas dadas de forma independente, as outras etapas também são descritas pelo autor.

É necessário que haja a coordenação de atividades para que a soma dos trabalhos individuais gere resultados de forma colaborativa. A coordenação consiste na divisão igualitária de tarefas entre os indivíduos e auxilia para que esforços de comunicação e cooperação não sejam perdidos.

O ato de cooperar reside na operação conjunta dos indivíduos de um grupo em um espaço compartilhado. Quando se trata de um espaço virtual, a cooperação dos indivíduos pode ser realizada por meio da produção, manipulação e organização de informações em instrumentos informatizados.

A captura de informações pode ser denotada como percepção, segundo Fuks (2001), a percepção é inerente ao ser humano, tornando-se o elemento central para a comunicação, coordenação e cooperação de um trabalho em grupo. A partir de ações os indivíduos tomam ciência das mudanças causadas nos ambientes que compartilham e podem redirecionar suas atitudes.

## 4.2 FERRAMENTAS COLABORATIVAS

As ferramentas de aprendizagem virtual colaborativa podem ser classificadas entre síncronas e assíncronas (RABELLO, 2010). As ferramentas síncronas necessitam que o usuário esteja conectado em um ambiente para que a interação seja realizada em tempo real. Exemplos de ferramentas síncronas são *chats*, videoconferências, ligações, *whatsapp*, entre outras. As ferramentas classificadas como assíncronas se dão pela interação entre indivíduos em qualquer lugar, não necessariamente em tempo real. Os fóruns, repositórios de arquivos e *wikis* são exemplos deste tipo de ferramenta.

Uma das ferramentas mais disseminada em ambientes de aprendizagem colaborativa e até mesmo em redes sociais é o *chat*, nele a interação entre usuários é realizada por meio do envio e recebimento de mensagens textuais que podem ser analisadas por frases já definidas em um sistema ou até mesmo por processamento de linguagem natural (PIETRUCHINSKI, 2016).

Os fóruns ou grupos de discussões são uma ferramenta que auxiliam na colaboração a partir de questões inseridas em um determinado ambiente e abertas a discussão para usuários. Nesta área um dos interlocutores apresenta a questão e os demais tem a oportunidade de discutir, apresentar sugestões, dúvidas e responder ou contradizer uma resposta (PIETRUCHINSKI, 2016). Os indivíduos normalmente se sentem mais à vontade para apresentar suas ideias nestas ferramentas do que em reuniões pessoalmente, sabendo que há um monitoramento de suas participações a qualidade das interações dos indivíduos pode se tornar maior do que em sistemas que não possuam monitoramento (EASTMAN; SWIFT, 2002).

Conceder notas as contribuições de indivíduos tanto em fóruns de discussão quanto em *chats*, se torna um grande desafio, como por exemplo, dar nota a um grupo de indivíduos para um projeto que necessita o ajuste de nota baseado no quanto um indivíduo contribuiu para a resolução do respectivo projeto ou classificar indivíduos perante a sua participação em um determinado curso (EASTMAN; SWIFT, 2002).

A criação de Ambientes Virtuais de Aprendizagem (AVAs), tem o intuito de auxiliar a análise de colaboração perante ferramentas do gênero. Este tipo de ambiente apresenta maneiras de monitorar a avaliação de atividades realizadas como: quantidade e conteúdo de postagens de indivíduos, data, hora e até quantidade de acessos nas ferramentas ou no ambiente. Porém, os administradores encontram

dificuldades para realizar o monitoramento e avaliação de contribuição dos indivíduos em um ambiente de aprendizagem virtual (LEE; TERASHIMA, 2012). Dessa forma, surgiu o *Collabora* como estudo de caso para este presente trabalho.

### 4.3 COLLABORA

O ambiente *Collabora* utiliza como base a ferramenta de colaboração *chat*, princípios da Aprendizagem Colaborativa Apoiada por Computador (CSCL) e Objetos de Aprendizagem (OA). A arquitetura do ambiente é capaz de promover e avaliar a colaboração com o intuito de obter dados que possam ser analisados da colaboração de um grupo ou estudante (ISHIKAWA *et al.*,2017). Um exemplo de atividade realizada é ilustrado na Figura 8.

**Figura 8 - Tela de atividade sendo realizada no *Collabora***

Fonte: Ishikawa *et al.* 2017

Conforme ilustrado na Figura 8, na plataforma de usuário o ambiente permite que as atividades possam ser iniciadas quando dois ou mais usuários estiverem *online*. Ao iniciar uma atividade, uma página com o enunciado do exercício escolhido é exibida com alternativas relacionadas à resolução da atividade e os usuários podem interagir por meio de um *chat*, para chegarem a um consenso e solucionar a atividade.

O ambiente ainda mantém um registro das interações realizadas pelos

usuários e exibe na plataforma do professor para que o mesmo realize a análise qualitativa das interações realizadas pelo grupo e pelos indivíduos de forma manual.

Na Figura 9 nota-se que o professor terá o trabalho de selecionar os *checkboxs* de inúmeras mensagens para que as mesmas sejam pontuadas como significativas ou não para o cálculo de colaboração. É neste momento que o mapeamento ontológico se torna importante para o estudo de caso *Collabora*.

**Figura 9 - Tela de análise de mensagens no *Collabora***

Dt. Postagem	Aluno	Mensagem	Limp	Significativa	Link	Imagem
28/03/2017 08:53	Rebeca Maria Garcez De Oliveira Dama	sim, araujo já tá vindo também	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28/03/2017 08:53	Maria Angela Domingues Da Silva	sim	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28/03/2017 08:53	Tamires Marques De Souza Rodrigues	todas aqui?	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28/03/2017 08:54	Maria Angela Domingues Da Silva	btz	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28/03/2017 08:54	Rebeca Maria Garcez De Oliveira Dama	a A está certa, acabei de ver na apostila	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28/03/2017 08:55	Tamires Cristina Araujo Cruz	entreei	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28/03/2017 08:56	Rebeca Maria Garcez De Oliveira Dama	beleza kk	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
28/03/2017 08:56	Tamires Cristina Araujo Cruz	a a ta certa mesmo	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Fonte: Ishikawa *et al.* (2017)

O desenvolvimento de atividades colaborativas permite a interação entre os participantes, porém a dificuldade está em avaliar o quanto cada aluno participou em sua finalização.

O *Collabora* contempla medidas de colaboração, como após a execução e uma atividade a plataforma soma o número total de colaboração dos membros de cada grupo em uma determinada atividade. O valor encontrado é utilizado como referência para o cálculo do *score* da colaboração de cada integrante de um grupo.

Algumas destas medidas adotadas no *Collabora* estão descritas no Quadro 4 e foram criadas com referências no trabalho de Ferraz *et al.* (2015), alguns critérios foram apresentados por Pietruchinski e Pimentel (2015), Khandaker e Soh (2010), Dascalu *et al.* (2010) e Yamada *et al.* (2016).

**Quadro 4 - Algumas medidas de avaliação utilizadas no *Collabora***

Quando avaliar a colaboração	Medidas	
<b>DURANTE</b> a execução de um <b>EXERCÍCIO</b> e <b>APÓS</b> a execução da <b>ATIVIDADE</b>	Integrante	Grupo
	Quantidade de mensagens postadas por integrante	Quantidade de mensagens postadas pelo grupo
	Quantidade de mensagens significativas postadas por integrante	Quantidade de mensagens significativas postadas por grupo
	Quantidade de mensagens de incentivo postadas pelo integrante	Quantidade de mensagens de incentivo postadas pelo grupo
	Quantidade de links e arquivos compartilhados pelo integrante	Quantidade de links e arquivos postados pelo grupo
	Quantidade de links e arquivos significativos compartilhados	Quantidade de links e arquivos compartilhados pelo grupo
	Estar conectado durante a realização do exercício	Quantidade de links e arquivos significativos compartilhados pelo grupo

Fonte: Autoria própria

A proposta de aplicação da ontologia para mapeamento das mensagens dos participantes durante a execução de uma atividade deve ser capaz de responder a quantidade de mensagens significativas postadas por integrante/grupo e a quantidade de mensagens de incentivo postadas pelo integrante/grupo, tendo em vista que a plataforma já consegue realizar de forma automática a avaliação da colaboração dos participantes.

A questão está em avaliar a significância das mensagens por meio de uma ontologia e automatizar o trabalho do professor que atualmente necessita avaliar mensagem por mensagem para classificar como significativa.

#### 4.4 CONSIDERAÇÕES DO CAPÍTULO

Este capítulo apresentou conceitos gerais e ferramentas relacionadas as atividades colaborativas e o ambiente de aprendizagem *Collabora*.

A aprendizagem colaborativa ajuda aumentar a produtividade de indivíduos e os ambientes proporcionam a criação de conhecimento por meio da interação entre os participantes.

Conforme apresentado o ambiente *Collabora* não consegue identificar as mensagens colaborativas e necessita que as mesmas sejam classificadas de forma manual. Por isto, este trabalho propõe o uso de mapeamento ontológico para determinar se uma mensagem trocada entre os participantes é significativa ou não. O mapeamento proposto é detalhado no próximo capítulo.



## 5 MAPEAMENTO ONTOLÓGICO PROPOSTO PARA CLASSIFICAÇÃO DE MENSAGENS

Este capítulo descreve o processo de construção do mapeamento por meio de ontologias. A proposta de ontologia foi realizada por meio de uma investigação do domínio de Ambientes Virtuais de Aprendizagem (AVA) com intuito de classificar diferentes tipos de mensagens dos usuários do ambiente *Collabora*. A construção da ontologia foi realizada usando a metodologia *Ontology Development 101*, exceto a etapa de Reutilização de Ontologias Existentes pois não foi encontrada nos repositórios disponíveis na literatura ontologias com domínio similares ao propósito deste trabalho. A implementação da ontologia foi realizada na ferramenta Protégé. A seção 5.1 relata o que foi realizado na etapa de Domínio e Escopo da Ontologia de Referência. A seção 5.2 descreve a etapa de Enumeração de Termos e Classes. A definição de propriedades é apresentada na seção 5.3. A seção 5.4 detalha a criação de instâncias. Por fim, a última seção descreve as considerações finais do capítulo.

### 5.1 DOMÍNIO E ESCOPO DA ONTOLOGIA DE REFERÊNCIA

Considerando as perguntas essenciais descritas na subseção 2.5.1 do capítulo 2, designa-se que a ontologia de referência para o mapeamento possui como domínio a participação dos alunos no sistema de aprendizagem colaborativa denominado *Collabora*.

O objetivo da ontologia proposta é a disposição de um vocabulário relacionado ao contexto da Estatística para cruzamento de informações com as ontologias que serão geradas por meio das perguntas inseridas no sistema e interações dos participantes por meio do *chat*. Quem utilizará esta ontologia é o sistema *Collabora* como base de conhecimento para avaliar a significância das mensagens trocadas pelo usuário.

Além disto, algumas perguntas mais específicas foram formuladas para a ontologia. Estas perguntas foram criadas a partir do estudo das necessidades do ambiente *Collabora*, estas foram abordadas no capítulo anterior. Dessa forma, a base de conhecimento da ontologia proposta deverá ser capaz de responder questões de competência:

- Quais mensagens foram significativas para a resolução do problema?

- Quais mensagens estão relacionadas com o enunciado?
- Quais mensagens estão relacionadas com a resposta?

## 5.2 ENUMERAÇÃO DE TERMOS E CLASSES

O terceiro passo na metodologia é responsável por listar todos os conceitos importantes relacionados ao domínio da ontologia. Dessa forma, os conceitos abstraídos para o domínio de Participação são baseados em trabalhos já existentes na literatura para a avaliação de tarefas que envolvem diálogo.

Neste trabalho o diálogo é feito por meio de *chat* no *Collabora*. Além dos trabalhos já existentes na literatura, houve uma análise e entendimento de exemplos da avaliação de mensagens realizadas na plataforma para extração de quais mensagens poderiam ser consideradas como significativas. Em um dos exemplos disponíveis para análise verificou-se que as mensagens significativas poderiam ser divididas entre dois tipos, conforme ilustrado na Figura 10.

Figura 10 - Exemplo de colaboração

Collabora						
03:26						
02/04/2017 03:26		283	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
02/04/2017 03:26		tendi	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
02/04/2017 03:26		fica mais rapido	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
02/04/2017 03:26		dos jutamos e dividimos	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
02/04/2017 03:26		tem que ser tudo junto	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
02/04/2017 03:26		como assim	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
02/04/2017 03:26		a eu faco dos q tem 2k	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
02/04/2017 03:27		2833,33	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
02/04/2017 03:27		no meu deu 3.400	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
02/04/2017 03:28		2 deu 3.4 e 2 deram 283	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
02/04/2017 03:28		dividi por 50	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
02/04/2017 03:28		oxi	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
02/04/2017 03:28		vo fazer de novo	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
02/04/2017 03:28		te errado	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
02/04/2017 03:28		o meu deu 3,4	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Fonte: Collabora

O primeiro tipo de mensagens classificadas como significativas na plataforma segundo Ishikawa *et al* (2017), são mensagens que possuem relação com o domínio da Estatística. Ou seja, mensagens diretamente relacionadas aos exercícios propostos aos alunos. Podendo ser mensagens com termos numerais ou termos relacionados à Estatística. No exemplo da Figura 10 se tem um termo numeral. Com isto, o parâmetro de correlação com o enunciado torna-se importante como um marcador para a análise de significância das mensagens.

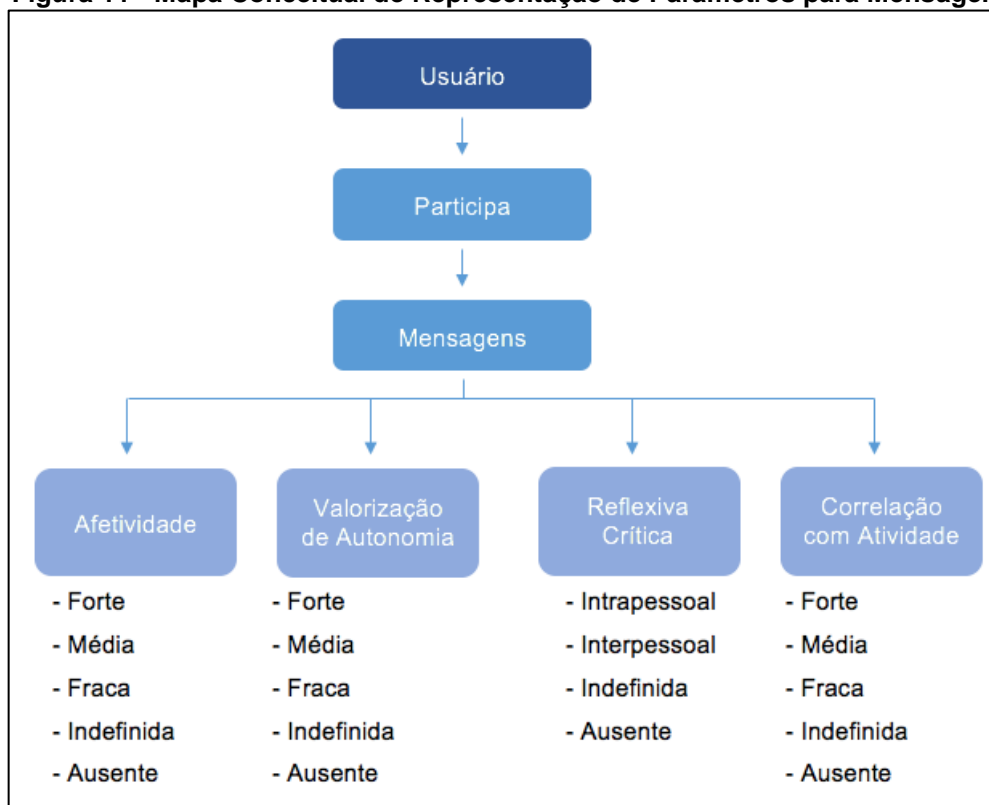
A segunda mensagem representa o segundo tipo de mensagens classificadas como significativas. São as mensagens que transmitem algum tipo de emoção que pode influenciar a participação e resolução dos exercícios. Com isto, alguns marcadores linguísticos podem ser utilizados para a classificação deste tipo de mensagens.

Para a identificação de quais marcadores linguísticos podem ser utilizados para a classificação deste tipo de mensagens utilizou-se do trabalho de David *et al* (2015), que consiste na análise do diálogo entre mensagens por meio da teoria de Freire (2006).

Os parâmetros utilizados para a classificação do diálogo entre participantes do *Collabora* foram testados em diversas plataformas e fóruns de participação. Os resultados obtidos foram essenciais para comprovar a fundamentação dos parâmetros escolhidos para a classificação.

Os pilares do diálogo foram divididos em parâmetros que se dão por marcadores linguísticos. Para este trabalho utiliza-se os parâmetros refinados e que se encaixam na proposta de aplicação da ontologia. Tais parâmetros são: (i) afetividade; (ii) valorização de autonomia; (iii) reflexividade crítica e sua relação com o enunciado pelo conceito; e (iv) correlação com enunciado. Um mapa conceitual com os tipos de mensagens possíveis e seus pesos é ilustrado na Figura 11.

**Figura 11 - Mapa Conceitual de Representação de Parâmetros para Mensagens**



**Fonte: Autoria própria**

De uma forma geral, pode-se dizer que ao analisar uma mensagem ela pode estar de acordo com algum dos parâmetros: afetividade, valorização da autonomia e correlação com a atividade, ilustrados na Figura 11 em sua forma forte, média, fraca, indefinida e ausente. Uma exceção se faz apenas para o parâmetro de reflexiva crítica, podendo assumir valores do tipo intrapessoal, interpessoal, indefinida e ausente.

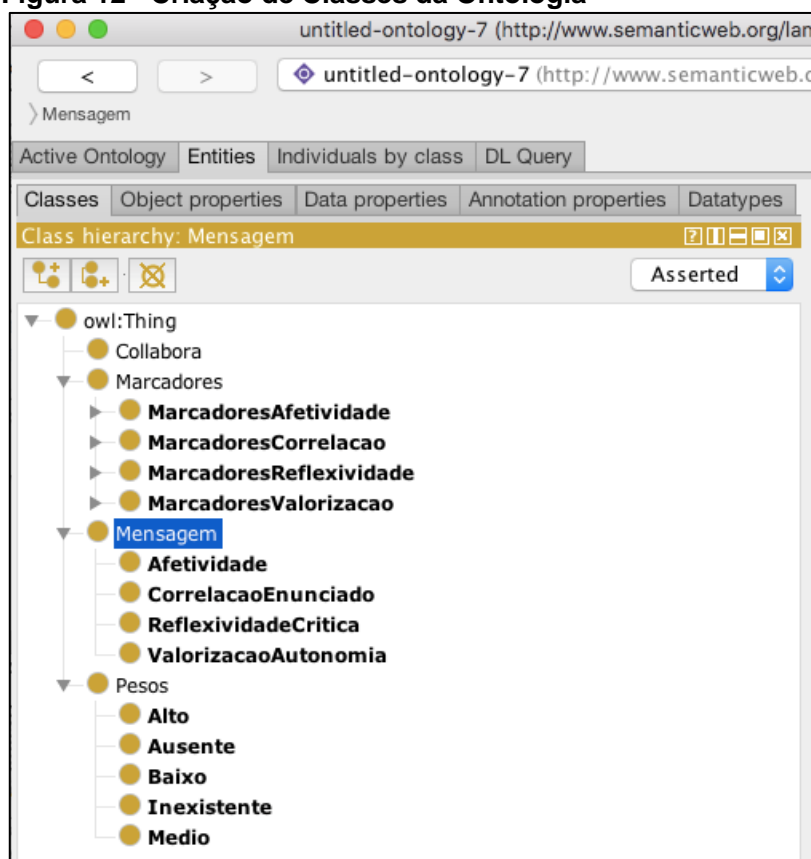
Sobre os valores forte, médio e fraco pode-se descrever o peso de cada mensagem. Por exemplo, ao classificar uma mensagem a partir da análise de seus termos e obter “muito bem” (referindo a elogio), “valeu” ou “obrigada” (referindo a gratidão), pode-se identificar a presença do parâmetro Afetividade, conforme relatado por David *et al.* (2015), com algumas modificações. Para isto, os marcadores são associados aos parâmetros que são verificados, tais como:

- Afetividade: elogios (“muito bem”, “parabéns”) concordância (“tem razão”, “está correto”), gratidão (“obrigada”, “valeu”), saudações (“olá”, “bom dia”), despedida (“até mais”, “tchau”), concessão (“desculpa”, “perdão”), entre outros.

- Valorização de Autonomia: estímulo (“isso aí”, “vá em frente”), sugestões (“eu acho que”, “sugiro que”), explicação (“por que?”, “me explica?”), entre outros.
- Reflexão Crítica: opiniões pessoais (“eu acho que”, “penso que”), afirmações (verbos no presente), exemplificação (“como por exemplo”), discordância (“não acha que”), entre outros.
- Correlação com Atividade: termos chave relacionados ao enunciado da atividade ou no domínio da estatística. Como por exemplo: média, mediana, valores, variância, teorema, cálculo e etc.

Com isto, cria-se as classes da ontologia referência. A ontologia criada possui as classes: Marcadores, Mensagem e Pesos. Além de uma classe denominada *Collabora*. Estas classes estão ilustradas na Figura 12, retirada da ferramenta *Protégé*.

**Figura 12 - Criação de Classes da Ontologia**



**Fonte: Autoria própria**

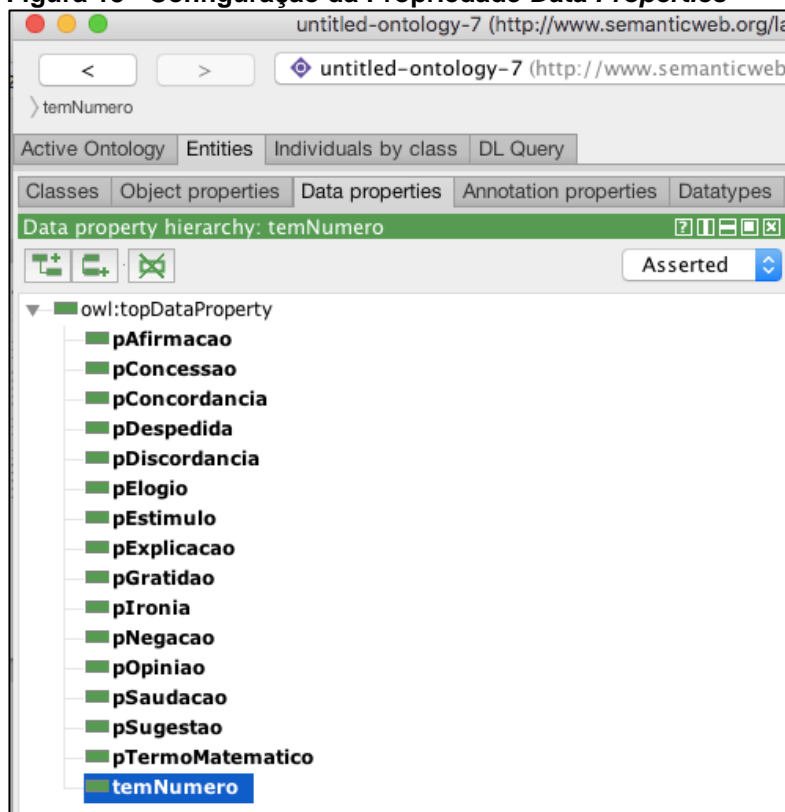
Observa-se que a classe Mensagem é responsável por conter os parâmetros especificados anteriormente como subclasses, tais como: afetividade, correlação com o enunciado, reflexividade crítica e valorização da autonomia. A classe de marcadores representa os marcadores linguísticos e de correlação que são utilizados para classificar as mensagens em diferentes parâmetros. A classe de pesos representa o grau de cada marcador, podendo ser do tipo alto, ausente, baixo, inexistente ou médio.

### 5.3 DEFINIÇÃO DAS PROPRIEDADES E RESTRIÇÕES

Apenas a criação de classes não fornece informações suficientes para responder as questões de competências que foram definidas anteriormente.

Dessa forma, foram definidos 16 marcadores linguísticos para se relacionar com a classe mensagem em *DataProperties*. Estes marcadores são os mesmos abordados anteriormente no qual as classes foram definidas. A Figura 13 ilustra a configuração destas propriedades.

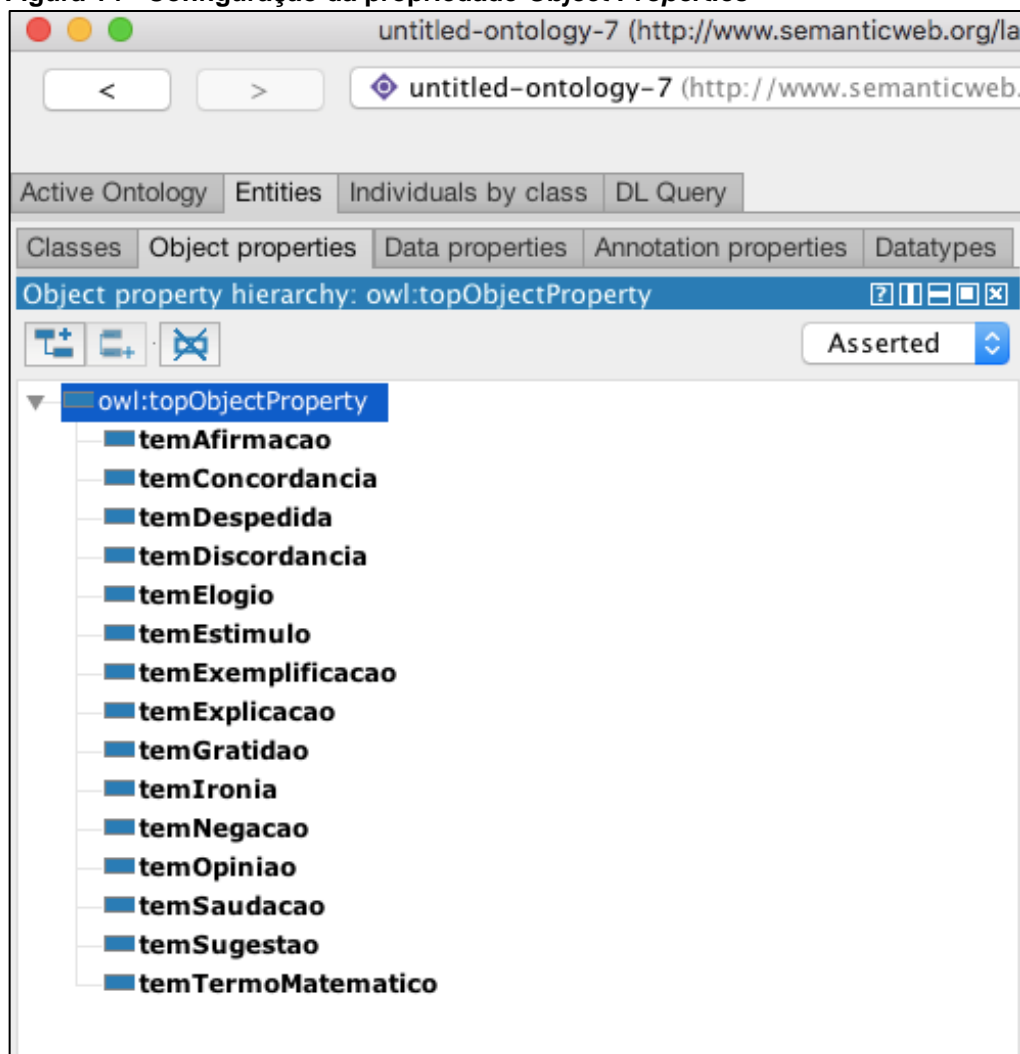
Figura 13 - Configuração da Propriedade *Data Properties*



Fonte: Autoria própria

Para as configurações das propriedades dos objetos (*Object Properties*) as configurações foram feitas para relacionamento com as subclasses de marcadores de acordo com a Figura 14.

Figura 14 - Configuração da propriedade *Object Properties*



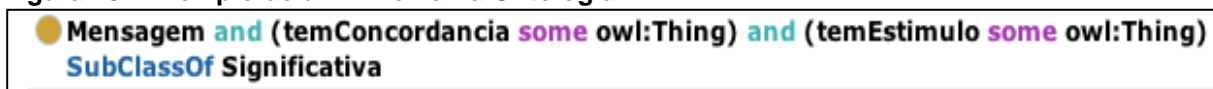
Fonte: Autoria própria

Estas propriedades representam as propriedades dos objetos que serão criados. Ou seja, caso possuam afirmação, concordância, termo matemática e entre outros.

O próximo passo é o de criação de axiomas, ou seja, regras para a classificação das mensagens nas diferentes classes e propriedades da ontologia. Sabe-se que uma mensagem poderá conter diferentes parâmetros. Desse modo, alguns *axiomas* foram gerados além das propriedades de relacionamento já

implementadas. Um exemplo de um *axioma* classificada como Significativa a uma mensagem que possui um nível de Concordância e Estimulo se encontra na Figura 15.

Figura 15 - Exemplo de um Axioma na Ontologia



Fonte: Autoria própria

Um total de 42 *axiomas* (regras), foram gerados neste trabalho para fazer restrições às regras inerentes as instâncias que foram criadas e testadas na ontologia. Do total de 42 *axiomas*, 16 foram criados com a função de apenas classificar o tipo de mensagem a classe que ela corresponde, como exemplificado na Figura 16.

Figura 16 - Exemplo de um axioma padrão



Fonte: Autoria própria

Além destes *axiomas*, outros 16 foram criados para classificar mensagens que já foram classificadas em alguma das classes da ontologia de referência, também classificadas como Significativas. A Figura 17 ilustra este tipo de *axioma*.

Figura 17 - Exemplo de um axioma para significativa



Fonte: Autoria própria

Por fim, 10 *axiomas*, como o ilustrado na Figura 15, foram criados com operações de *and* (E) ou *or* (OU) para auxílio da classificação das instâncias. A próxima seção abordará a criação das instâncias para aplicação deste trabalho.

## 5.4 CRIAÇÃO DAS INSTÂNCIAS

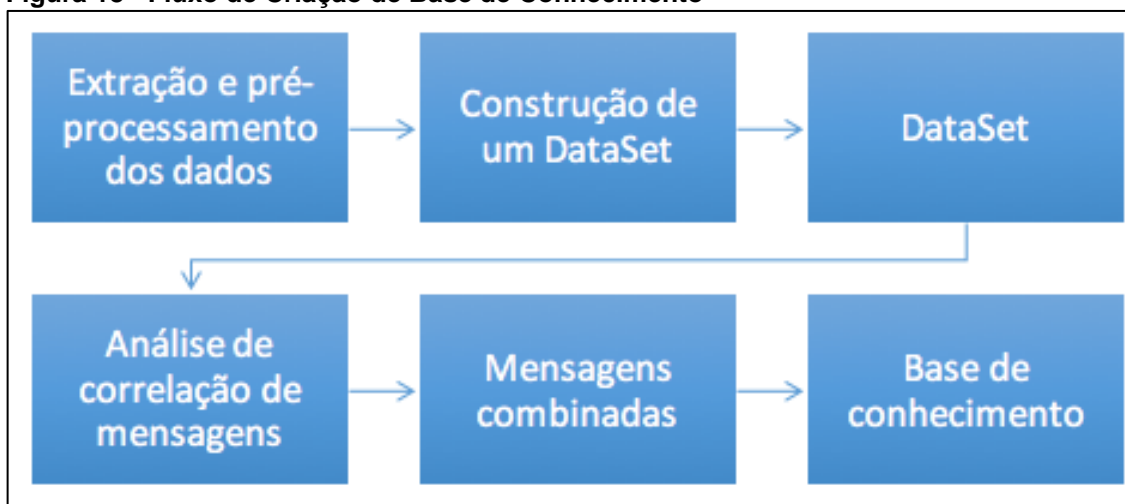
O último passo é o de criação de instâncias para as classes. Um conjunto de instância é responsável por constituir a base de conhecimento da ontologia proposta.



Um fluxo de criação de base de dados é proposto com base no funcionamento de ambientes de aprendizagem colaborativa como o *Collabora*. Neste caso, após a colaboração realizada dentro de um ambiente, é necessário a extração e pré-processamento dos dados obtidos através do *chat*. Com estes dados obtidos inicia-se o processo de criação de um *dataset* para realizar a aplicação da proposta do trabalho que é a de analisar a correlação das mensagens com os marcadores inseridos na ontologia e encontrar mensagens para construir uma base de conhecimento que pode ser utilizada e reutilizada no processo de análise de mensagens significativas.

Considerando o desenvolvimento da base da ontologia proposta sugere-se divisão de etapas, conforma ilustra a Figura 18.

**Figura 18 - Fluxo de Criação de Base de Conhecimento**



Fonte: Autoria própria

As etapas da Figura 18 representam o fluxo de criação de base de conhecimento para a implementação do mapeamento por meio de ontologias no sistema.

A extração das mensagens já ocorre na plataforma, de forma que as interações dos usuários são extraídas do banco de dados e dispostas na Tela de Análise de Mensagens do Professor. É necessário um pré-processamento destes dados para a construção de um *DataSet*.

Com um *DataSet* construído é possível utilizar técnicas como a EAT (Extração Automática de Termos) baseada na EI (Extração de Informação). Esta técnica se trata do processo de identificar automaticamente tipos específicos de entidades, conceitos

e relações em textos livres e armazenar a informação processada de forma estruturada (YANGARBER; GRISHMAN, 2000).

Com a extração das mensagens e análise de correlação das mesmas por meio da ontologia de referência criada é possível verificar mensagens combinadas e construir uma base de conhecimento sólida para verificar a colaboração dos participantes por meio da significância das mensagens.

## 5.5 CONSIDERAÇÕES DO CAPÍTULO

Este capítulo mostrou a aplicação da metodologia *Ontology Development* no desenvolvimento de uma ontologia para determinar mensagens significativas trocadas durante a solução de atividades colaborativas do *Collabora*. Foram criadas 16 propriedades e 42 axiomas.

Os resultados da ontologia proposta estão descritos no próximo capítulo.

## 6 RESULTADOS

Este capítulo descreve os resultados obtidos por meio de testes da ontologia criada aplicada a base de dados obtida pelo *Collabora*. A seção 6.1 relata a classificação de mensagens que possuem correlação direta com atividades propostas no *Collabora*. A seção 6.2 apresenta a classificação de mensagens que possuam marcadores linguísticos também utilizados como parâmetros para classificação de mensagens significativas. Por fim, a última seção deste capítulo descreve as considerações finais do capítulo.

### 6.1 MENSAGENS COM CORRELAÇÃO COM ATIVIDADE

O banco de dados para a realização dos casos de teste foi obtido por meio ambiente *Collabora*. O Quadro 5 ilustra a ementa com seus conteúdos dispostos de acordo com o banco de dados.

**Quadro 5 - Conteúdo de questões do *dataset* do *Collabora***

<b>Ementa</b>	<b>Conteúdo</b>
<b>Introdução à Estatística</b>	Visão Geral
	Tipos de Dados
	Pensamento Crítico
	Planejamento da pesquisa
<b>Estatística para descrição, exploração e comparação de dados</b>	Visão Geral
	Medidas de Tendência Central
	Medidas de Dispersão de Variação
	Medidas de Posição
<b>Estimação</b>	Estimativa de uma Média Populacional
	Estimativa de uma Proporção Populacional
	Estimativa de uma Variância Populacional
<b>Testes de Hipótese</b>	Teste de Afirmativa sobre uma média
	Teste de Afirmativa sobre uma proporção
	Teste de Afirmativa sobre desvio-padrão
	Comparação de duas variáveis
	Inferência sobre duas médias
	Comparação de variação de duas proporções

Fonte: Adaptado de *Collabora*

Os assuntos contemplados no *Collabora* abordam a Estatística e sua ementa contém: introdução à estatística com quatro conteúdos, estatística para descrição, exploração e comparação de dados com quatro conteúdos, estimação com três conteúdos e testes de hipótese abordando seis conteúdos. Os assuntos totalizam 17 tipos de conteúdos que podem ser inseridos em atividades do ambiente para a resolução.

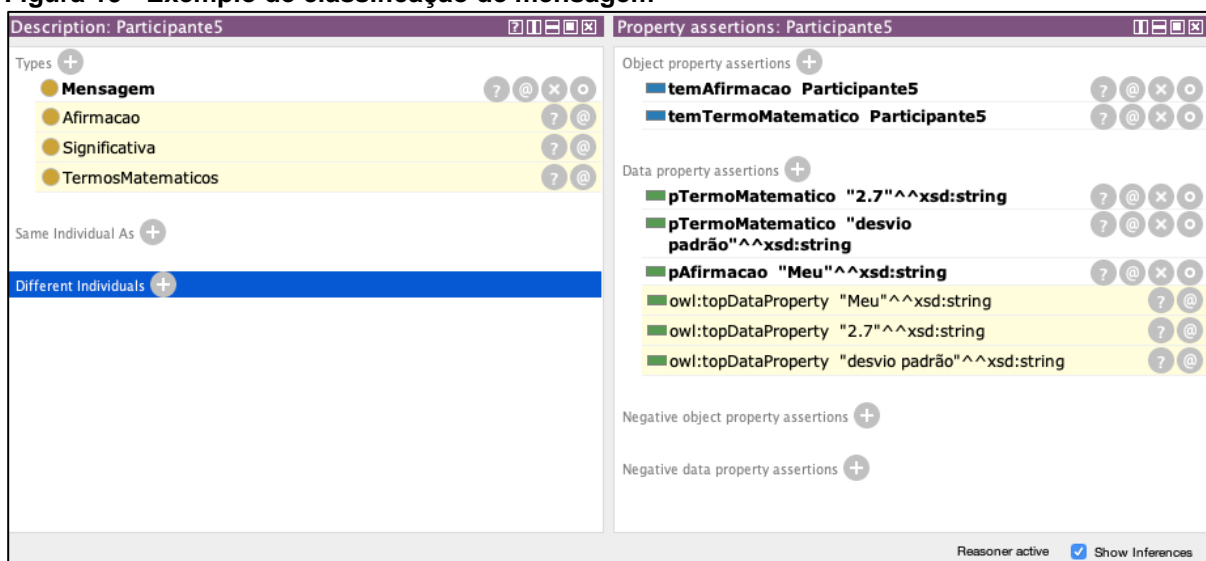
Uma atividade foi selecionada da base de dados para a realização de testes com a ontologia criada no capítulo anterior. A atividade escolhida está relacionada a ementa de estatística para descrição, exploração e comparação de dados e possui o conteúdo de visão geral e medidas de tendência central. A escolha desta atividade foi porque estava descrita na tese de Ishikawa (2018) e o *Collabora* não se encontrava acessível no momento de realização do experimento. A atividade selecionada possuía um total de 98 interações de usuários. As interações entre os usuários para esta atividade encontram-se no Anexo A.

Para tornar possível a realização da classificação das mensagens alguns termos estatísticos foram extraídos de repositórios da literatura relacionados ao domínio da Estatística para cruzamento entre as interações dos participantes das atividades. Estes termos estão descritos no Anexo B.

Com relação aos termos do domínio da Estatística e interações dos participantes um teste de cruzamento de mensagens e termos-chave foi realizado por meio da ontologia de referência proposta no capítulo anterior.

Após a realização do teste a ontologia de referência retornou um total de 33 mensagens com correlação ao domínio da Estatística de um total de 98 mensagens e que podem ser classificadas como mensagens significativa. A Figura 19 ilustra o exemplo de uma classificação de interação de usuário na ferramenta *Protégé*. A mensagem enviada pelo usuário foi: “Meu desvio padrão deu 2.7”

**Figura 19 - Exemplo de classificação de mensagem**



**Fonte: Autoria própria**

Neste exemplo de classificação é possível verificar que a ontologia de referência foi capaz de classificar a mensagem em diferentes marcadores associados as propriedades das mesmas. A ontologia classifica a mensagem contendo TermosMatematicos, com Afirmacao (que será abordada da próxima seção) e por conter estes marcadores a classifica como Significativa.

Todas as mensagens classificadas como significativas neste teste podem ser encontradas no Apêndice A. A próxima seção é responsável por apresentar os outros marcadores inseridos na ontologia de referência da aplicação deste trabalho.

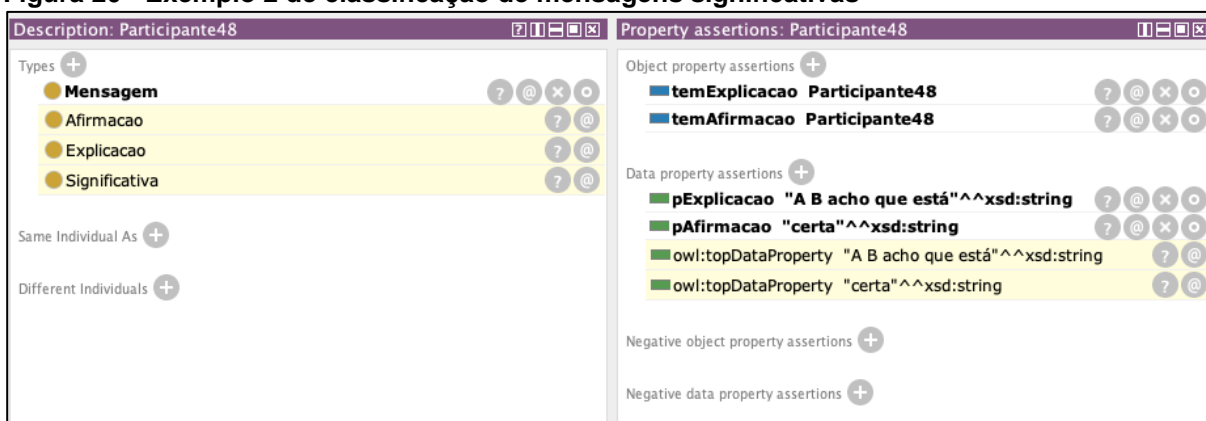
## 6.2 MENSAGENS COM MARCADORES LINGUÍSTICOS

Os marcadores linguísticos selecionados para a aplicação da ontologia proposta foram documentados no capítulo anterior. Esta seção é responsável por ilustrar um teste realizado com uma atividade classificando mensagens com estes marcadores. A atividade escolhida foi a ATV\_01\_07 - AT2 que pode ser encontrada no Anexo C, contém um total de 90 interações.

Após a realização do teste a ontologia de referência retornou um total de 23 mensagens com marcadores linguísticos e que podem ser classificadas como mensagens significativas.

A Figura 20 ilustra o exemplo de uma classificação de interação de usuário na ferramenta *Protégé*. A mensagem enviada pelo usuário foi: “a B acho que está certa”

**Figura 20 - Exemplo 2 de classificação de mensagens significativas**



**Fonte: Autoria própria**

Neste exemplo de classificação verifica-se que a ontologia de referência classificou a mensagem em diferentes marcadores linguísticos e significativa. Classifica a mensagem contendo Afirmação “eu acho” “certa”, Explicação “eu acho” e por conter estes marcadores a classifica como mensagem Significativa.

Todas as mensagens classificadas como significativas neste teste realizado podem ser encontradas no Apêndice B.

### 6.3 BREVE ANÁLISE DO MAPEAMENTO PROPOSTO

A avaliação da colaboração utilizando o mapeamento ontológico proposto pode ser realizada em ambientes colaborativos, considerando os seguintes indicadores:

- Número de mensagens significativas alinhadas;
- Número de mensagens trocadas durante a realização de uma atividade;
- Número de parâmetros da ontologia da ontologia de referência.

Algumas métricas podem ser geradas como:

- Número de mensagens significativas de um participante em relação a todos os participantes;
- Número médio de mensagens significativas por participante.

Portanto, o mapeamento ontológico contribui para a avaliação das mensagens significativas de forma mais automatizada. Uma proposta futura é que o mapeamento

proposto no *Protégé* possa ser integrado ao *Collabora* deixando o trabalho mais automatizado.

#### 6.4 CONSIDERAÇÕES DO CAPÍTULO

Este capítulo mostrou os resultados da aplicação do mapeamento proposto usando como base de dados as interações entre os participantes durante a resolução das atividades colaborativas presentes no ambiente *Collabora*.

Com isto foi possível observar que a ontologia de referência criada para aplicação é capaz de classificar e identificar diferentes tipos de mensagens em diversos marcadores e ainda classificar se são significativas para o processo de colaboração do ambiente.

## 7 CONCLUSÃO

Este trabalho propôs uma ontologia de referência a fim de verificar se mensagens trocadas entre participantes em um ambiente colaborativo são significativas ou não em relação as atividades. A ontologia proposta contempla parâmetros linguísticos para auxiliar a análise e classificação das mensagens significativas.

Foi realizada uma revisão da literatura sobre ontologias, trabalhos relacionados e metodologias. A metodologia escolhida para o desenvolvimento da ontologia foi a *Ontology Development 101* dentre as citadas neste trabalho.

Destaca-se a comparação entre ferramentas encontradas na literatura antes da escolha e utilização do editor de ontologias *Protégé*. Com o uso da ferramenta *Protégé* apoiada pela sua metodologia sugerida, e da análise do funcionamento do estudo de caso para este trabalho, o ambiente *Collabora*, foi possível identificar os aspectos que podem ser considerados como colaboração e, com isso, fundamentar e documentar uma ontologia de aplicação para classificação de mensagens significativas em ambientes de colaboração.

O trabalho não pôde realizar testes reais na plataforma *Collabora*, porém conseguiu desenvolver cenários ilustrativos com bases de dados reais da própria plataforma para a avaliação de mensagens e validação deste projeto. As bases de dados foram extraídas diretamente do banco de dados do ambiente.

Vale ressaltar que embora a ontologia de referência criada não seja um padrão inquestionável de avaliação, se a mesma for refinada e mais fundamentada no domínio da Estatística poderá configurar-se como uma boa referência de comparação para avaliação da aprendizagem em ambientes de colaboração deste mesmo domínio.

### 7.1 TRABALHOS FUTUROS

Como trabalhos futuros desta pesquisa sugere-se:

- Aplicar a proposta de ontologia criada no ambiente *Collabora* para a realização de testes reais dentro do ambiente.



- Analisar qual é a melhor técnica de extração das mensagens para construção de base de conhecimento por meio do processamento de mensagens obtidas.
- Refinar os *axiomas* da ontologia para abranger mais relações e restrições entre as instâncias e classes da ontologia.
- Comparar os resultados obtidos por meio do uso da ontologia proposta com os resultados do ambiente atual.
- Ampliar a quantidade de palavras-chaves utilizadas na ontologia proposta.

## REFERÊNCIAS

- ABBAS, M.; BERIO, G. Creating ontologies using ontology mappings: Compatible and incompatible ontology mappings. In: IEE/WIC/ACM International conference on web intelligence (WI) and intelligent agent technology. **Proceedings...** Atlanta, 2013, p. 143-146.
- ABOLHAHASSANI, H.; *et al.* **On Ontology Alignment Experiments**. Disponível em: <<http://www.webology.org/2006/v3n3/a28.html>> Acesso em 16 de setembro de 2017.
- ALATRHS, E. Comparison of Ontology Editors. **eRAF Journal on Computing**, v. 4, n.1, p.23-38, 2012.
- BARRO, S B.; MIZOGUICHI, R.; VERDEJO, F. A Platform for Collaboration Analysis in CSCL: An ontological Approach. In: AIED'2001. **Proceedings...** Texas, 2001. p.19 - 25.
- CESARE, S; GEERTS, L; HOLLAND, G; LYCETT, M; PARTRIDGE, C. Ontology-driven software engineering. In: 24th ACM Sigplan Conference on Object Programming Systems Languages and Applications. **Proceedings...** Orlando, 2009, p-723-724.
- CHOI, N. A Survey on Ontology Mapping. *ACM SIGMOD Record*, v. 35, n.3, p.34-41, 2006.
- COPPIN, Ben. **Artificial Intelligence Illuminated**. 1 ed. LTC, 2004.
- CORSKUN, E; GRABOWSKI, M. Software complexity and its impacts in embedded intelligent real-time systems”, **Journal of Systems and Software**, v.78, n.2, p.128-145, jan, 2005.
- DASCALU, M.; TRAUSAN-MATU, S.; DESSUS, P. Utterances Assessment in chat conversations. **Research in Computing Science**, v.46, n.1 p.323-334, 2010.
- DEGEN, W., HELLER, B., HERRE, H., SMITH, B. Gol: “Toward an axiomatized upper-level ontology.” In: ACM New York International Conference on Formal Ontology in Information Systems. Ogunquit, Maine, **Proceedings...** USA, 2001, p.34-46.
- EASTMAN, J; SWIFT, C; E. Enhancing collaborative learning: Discussion boards and chat rooms as project communication tools. **Business Communication Quarterly**. v. 65, n.3, p. 29-41, 2002.

FALBO, R.; *et al.* Towards Semantic Software Engineering Environments. In: 14<sup>th</sup> International Conference on Software Engineering, Ischia, Itália, **Proceedings...** Itália, 2002, p. 477-478.

FAYYAD, U; PIATETSKY-SHAPIRO, G; SMYTH, P. From Data Mining to Knowledge Discovery in Databases. **American Association for Artificial Intelligence**. v.17, n.3, p. 45-49, mar, 1996.

FENSEL, D.; HERMELEN, F. van. On-To-Knowledge: Content-Driven Knowledge Management Tools through Evolving Ontologies. Disponível em: <http://www.ontoknowledge.org/down/del33.pdf>. Acesso em 13/04/2018.

FERRAZ, P O.; DE OLIVEIRA, P.; HORNINK, G.G. Desenvolvimento e Implementação de Indicadores de Colaboração e Participação no Moodle. **Informática na educação: teoria & prática**, v. 18, n. 1, p. 85 - 95, 2015.

FUKS, H.; *et al.* Apllying the 3C Model To Groupware Develipment. **International Journal of Cooperative Information System**. v.14, n.1, p. 299-328, 2001.

FUSSEL, R.; KREUZ, J. Social and Cognitive Approaches to Interpersonal Communication: Introduction and overview. **Social and Cognitive Approaches**. v.12, n.1, p. 73-79, 1998.

GRUBER, T. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. In: International Journal-Human-Computer Studies. Padova, Italy, **Proceedings...** Itália, 1993, p. 907-928.

GUARINO, N; GIARETTA, P. Ontologies and Knowledge Bases: Towards a terminological clarification. **Towards Very Large Knowledge Bases**. v.1, n.1, p.25-32, 1998.

GUARINO, N. Formal Ontology and Information Systems, In International Conference on Formal Ontology in Information Systems, Trento, Italy, **Proceedings...** Trento, Italy, 1998, p.3-15.

GUARINO, N. e WELTY, C. Evaluating Ontological Decisions With Ontoclean. **Communications of the ACM**, v. 45, n. 2, p. 61–65, 2002.

GU, J. C. H; YANG, L; ZHANG, L. OBSA: Ontology-based semantic information processing architecture. In: IEE/WIC/ACM International Conference On Web Intelligence. **Proceedings...** Washington, 2004, p. 607-610.

GUIZZARDI, G. **Ontological foundations for structural conceptual models**. Holanda, 2005. 416p. Tese de Doutorado — Centre for Telematics and Information Technology, University of Twente.

GUIZZARD, G.; WAGNER, G. Towards Ontological Foundations for Agent Modeling Concepts Using UFO. In: Agent-Oriented Information Systems II, Heidelberg, Berlin, **Proceedings...** Heidelberg, Berlin, 2005, p. 110-124.

HAGUENAUER, C *et al.* Ambientes colaborativos de aprendizagem no apoio ao ensino presencial: a experiência do programa de pós-graduação. **Colabor@-A Revista Digital da CVA-RICESU**, v. 4, n. 16, 2010.

ISHIKAWA, E; BELUZZO, L; MATOS, S; SANTOS, G. Collabora: A Collaborative Architecture for Evaluating Individuals Participation During the Development of Activities. **International journal of software engineering & applications (IJSEA)**, v. 8, N.1, p. 33-48, 2017.

ISHIKAWA, E. **Objeto Virtual de Aprendizagem Colaborativa (Collabora): Estudo na disciplina de probabilidade e estatística no ensino superior**. Ponta Grossa, 2018. 200p. Tese de Doutorado – Programa de Pós-Graduação em Ensino de Ciência e Tecnologia, Universidade Tecnológica Federal do Paraná.

ISOTANI, S; BITTENCOURT, I; BARBOSA, E; DERMEVAL, D; PAIVA, R. Ontology driven software engineering: a review of challenges and opportunities. **IEEE Latin transactions**. v. 13, n. 3, mar 2015.

KALFOGLOU, Y.; SCHORLEMMER, M. IF-Map: An Ontology-Mapping Method Based on Information-Flow Theory. **Journal on Data Semantics**, v.1, n.10, p-78-94, 2003.

KHANDAKER, N.; SOH, K. Classroomwiki: A collaborative wiki for instructional use with multiagent group formation. **IEEE Transactions**, v.3, n.3, p.190-202, 2010.

LEE, Y; TERASHIMA, N; A distance instructional system with learning performance evaluation mechanism: Moodle-based educational system design. **International Journal of Distance Education Technologies (IJDET)**, v.10, n.2, p.57-64, 2012.

LMAT, I; *et al.* Alignment between two domains ontologies: Case of educational orientation in mathematics education. In: 5<sup>th</sup> Informational Conference on Principles and Practice of Declarative Programming. **Proceedings...** Morocco, 2015. p. 1-2.

LEIDIG, T. L3 Towards and Open Learning Environment. **ACM Journal of Educational Resources in Computing**. v.1, n.1 p.87-96, 2001.

LIDDY, R. **Natural Language Processing**, 2 ed. Library and Information Science, 2003.

NOY, F.; MCGUINNESS, L. **Ontology Development 101: A guide to creating your first ontology**. Disponível em: <  
<http://www.ksl.stanford.edu/people/dlm/papers/ontology101/ontology101-noy-mcguinness.html>> Acesso em 10 de fevereiro de 2018.

PIETRUCHINSKI, M. **ArqMAEC: Um modelo arquitetural baseado em Agentes para monitorar, avaliar e estimular a colaboração em ambientes educacionais gamificados**. Curitiba, 2016. 361p. Tese de Doutorado – Programa de Pós-Graduação em Informática, Universidade Federal do Paraná.

PIETRUCHINSKI, M.; PIMENTEL, R. An architectural model of multi-agent systems for student evaluation in collaborative game software. In: ICAISC 2015: International Conference on Artificial Intelligence and Soft Computing, Berlin, Germany, **Proceedings...** Berlin, 2015, p. 1542.

SILVA, V. **Uma Linguagem de Modelagem para Sistemas Multi-agentes Baseada em um Framework Conceitual para Agentes e Objetos**. Rio de Janeiro, 2004. 252p. Tese de Doutorado - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

RABELLO, R. **Interação e Autismo: uso de agentes inteligentes para detectar déficits de comunicação em ambientes síncronos**. Porto Alegre, 2010, 267p. Tese de Doutorado – Programa de Pós-Graduação em Informática na Educação, Universidade Federal DO Rio Grande do Sul.

ROCHA, F. Avaliação de aprendizagem: **Uma abordagem qualitativa baseada em mapas conceituais, ontologias e algoritmos genéticos**. Belém, 2007. 163p. Tese de Doutorado – Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal Do Pará.

ROCHA, F; COSTA, J; FAVERO L. Como usar ontologias na avaliação da aprendizagem significativa mediada por mapas conceituais. **Revista Brasileira de Informática na Educação**, v. 13, n. 1, p.53-64, 2005.

RUSSELL, Stuart; NORVIG, Peter. **A modern approach: Artificial Intelligence**. Prentice-Hall: Egnlewood Cliffs, 1995, v. 25, p. 27.

SHI, F; LI, J; TANG, G; XIE, T; LI, H. Actively Learning Ontology Matching via User Interaction. In: 8<sup>th</sup> International Web Conference on Principles and Practice Declarative Programming. **Proceedings...** Washington, 2009. p. 585-600.

SHVAIKO, P; EUZENAT, J. Ontology matching: state of the art and future challenges. **IEEE Transactions on Knowledge and Data Engineering**. v.25, n.1, p.158-176, nov 2013.

STAHL, G; KOSCHMAN, T; SUTHERS, D. Computer-supported collaborative learning: An historical perspective. **Cambridge handbook of learning sciences**, 2006.

WIMALASURIYA, D; DOU, D. Using multiple ontologies in information extraction. In: 18<sup>th</sup> ACM conference on information and knowledge. **Proceedings...** Hong Kong,

2009, p. 235-244.

WONGTHONGTHAM, P; CHANG, E; DILLON, T; SOMMERVILLE, I. Development of a Software Engineering Ontology for Multisite Software Development", **IEEE Transactions on Knowledge and Data Engineering**. v.21, n.8, p. 1205-1217, out 2009.

WONG, W; BENNAMOUN, M. Ontology learning from text: A look back and into the future. **ACM Computing Surveys (CSUR)**, v. 44, n.4, p. 82-117, ago 2012.

**ANEXO A - Mensagens retiradas de ATV\_01\_07 – AT3**

ALUNO	MENSAGEM
47	Pronto
47	O
48	Ooo
50	O/
50	Falta o Renan
56	to aqui gente
48	Não altera ?
48	Pelo celular é complicado mano
50	Creio que não
47	Bem complicado
47	Qual a fórmula?
48	Ou é A ou E
48	Ou ou ouuuu
47	Qual o valor inicial da variância?
47	Não lembro o que é
50	Nós não temos acesso à cada idade dos funcionários
47	A variância inicial é 4 certo?
47	As idades estão entre 28 e 36
50	O desvio padrão é de quatro
47	Não vai se alterar
47	Se a variação está entre esses valores
47	Se a outra idade for 32 não se altera
48	E então ?
50	Aí vc tá falando de desvio padrão ou de variância?
47	Só alteraria de o valor fosse maior que 36 ou menor que 28
47	Variância
56	pqp eu tava fazendo o calculo do desvio padrao
56	ah, eh so elevar ao quadrado?
50	Pra mim também não se altera, mas tem alguma coisa aí
47	Dado um conjunto de dados, a variância é uma medida de dispersão que mostra o quão distante cada valor desse conjunto está do valor central (médio).
47	Simples
47	Aaa
48	Mas então ta estranho mesmo, pq a distribuição de salários altera sim, se o mesmo valor for dividido por mais 2 funcionários é pra alterar
48	Teoricamente
47	Porém o que irá se alterar é a média
56	gente... o meu deu 5,42% a menos
56	O-O
48	O minha primeira conta deu 4,8%
47	Acho que é 5 mesmo
50	Como foi o cálculo que vcs fizeram?



47	Marquei a
47	Aquela fórmula que vc passou
48	Marquei a
56	agora deu exato
47	Confia no Irineu
56	16 - 100%
56	15,2 - x
56	x = 95%
56	entao o novo resultado seria 95% ou seja 5% menor
47	Bora letra a
56	o que eu fiz faz sentido?
56	nao sei
56	mas foi o que deu aqui
48	Faz
47	Si fouda
48	A media não altera mas a variante do salario sim, precisaria aumentar proporcionalmente o salario
56	eu respondi letra A
47	Bora
48	Pra não ter alteração
50	Onde vc pegou aquele 15,2?
47	Vai Rodrigues
56	Ja mando a foto do que eu fiz
48	Essa questão era meio logica rodrigues,
56	num era nao viado ahsufhasuehfuase
47	Eu achei
48	Era vey pensa comigo, se aumentar o numero de funcionarios, não ia alterar a media de idade, mas altera o valor de distribuição de salario
56	sim.. o problema era achar qto
56	enfim... bora questao 2
47	A Á lógica era que o valor iria mudar
48	38 = 100% -> logo aumento de 2 membros é = 4,8% de diferença de distribuição de salario kkk fiz de cabeça mas não façam isso em casa
47	Por ser um valor com com uma diferença pequena
47	O 5% seria justo
50	Blz
50	To tentando calcular esse daqui agora
56	a media é quanto?
56	pfvr
50	12,54
48	12,54
47	Meu desvio padrão deu 2.7
47	Mintiiira
48	o meu deu 1,759

47	A variação deu isso
50	O meu deu 1,76
56	o meu deu 1,76 tbm
48	variancia deu 3,096
47	Isso
47	Aqui também
47	C e D
48	C e D
50	Tb
50	3,09598
56	mas gente.. excel eh tao lindo ne
56	ele so resolve
47	Voa pai
50	C e D
50	Fechou?
47	Já respondi
56	tbm

**ANEXO B - Termos-Chave relacionados ao domínio da Estatística**

**Amostra** - é um subconjunto finito da população que se supõe representativo desta.

**Amostra Amodal** - é uma amostra que não tem moda.

**Amostra Bimodal** - é uma amostra que tem duas modas.

**Amostra Imparcial** - é uma amostra em que todos os elementos tiveram uma igual oportunidade de fazer parte da mesma.

**Amostra Multimodal** - é uma amostra que tem mais do que duas modas.

**Amostra Representativa** - é aquela que deve conter em proporção todas as características qualitativas e quantitativas da população.

**Amostragem Aleatória Simples** - é aquela em que qualquer elemento da população tem a mesma probabilidade de ser escolhido.

**Amostragem Estratificada** - é aquela em que a população está dividida em estratos ou grupos diferenciados.

**Amostragem Sistemática** - é aquela em que os elementos são escolhidos a partir de uma regra previamente estabelecida.

**Amplitude de um Conjunto de Dados** - é a diferença entre o maior valor e o menor valor desse conjunto. Se os dados estiverem agrupados em classes, a amplitude é a diferença entre o limite superior da última classe e o limite inferior da primeira.

**Atributos Qualitativos** - são atributos que estão relacionados com uma qualidade e apresentam-se com várias modalidades.

**Atributos Quantitativos** - são atributos aos quais é possível atribuir uma medida e apresentam-se com diferentes intensidades ou valores.

**Censo** - é um estudo estatístico que resulta da observação de todos os indivíduos da população relativamente a diferentes atributos pré-definidos.

**Classe Mediana** - é a classe, para dados classificados, que contem a Mediana (neste caso considera-se como Mediana o valor da variável estatística que corresponde a  $n/2$ , quer  $n$  seja par, quer  $n$  seja ímpar).

**Classe Modal** - é a classe, para dados classificados, que aparece com maior frequência.

**Coefficiente de Correlação Linear ( r )** - medida estatística que permite calcular o valor numérico correspondente ao grau de dependência entre duas variáveis, o qual varia entre -1 e 1.

**Correlação** - é a relação ou dependência entre as duas variáveis de uma distribuição bidimensional.

**Correlação Fraca ou Nula** - quando o Diagrama de Dispersão não permite o ajustamento de nenhuma reta, o que significa que  $r = 0$ . Diz-se, então, que não existe nenhuma relação entre as variáveis da Distribuição Bidimensional.

**Correlação Negativa Forte** - quando a reta de regressão, obtida a partir do Diagrama de Dispersão, tem declive negativo. A correlação é negativa quando  $r$  varia entre -1 e 0 e será tanto mais forte quando  $r$  se aproxima de -1.

**Correlação Negativa Perfeita ou Linear** - quando a reta de regressão, obtida a partir do Diagrama de Dispersão, tem declive negativo com  $r = -1$ .

**Correlação Positiva Forte** - quando a reta de regressão, obtida a partir do Diagrama de Dispersão, tem declive positivo. A correlação é positiva quando  $r$  varia entre 0 e 1 e será tanto mais forte quanto  $r$  se aproxima de 1.

**Correlação Positiva Perfeita ou Linear** - quando a reta de regressão, obtida a partir do Diagrama de Dispersão, tem declive positivo com  $r = 1$ .

**Dados Classificados** - são valores que uma dada variável pode tomar dentro de certo intervalo. Estes dados são classificados ou agrupados em classes.

**Dado Estatístico** - é o resultado da observação de um atributo/variável qualitativa ou quantitativa.

**Dados Simples** - são valores associados a uma dada variável e cuja representação é feita

através de uma tabela.

**Definição do Problema** - é a primeira fase do estudo estatístico e consiste na definição e formulação correta do problema a ser estudado.

**Desvio Médio (  $\bar{d}$  )** - é a média aritmética do valor absoluto da diferença entre cada valor e a média, no caso dos dados não classificados. No caso dos dados classificados, tem que se entrar em conta com a frequência absoluta de cada observação.

**Desvio Padrão** - é a raiz quadrada positiva da variância.

**Diagrama de Dispersão** - é a representação num referencial ortonormado de um conjunto de pares ordenados de valores  $(x, y)$ , onde cada par ordenado corresponde a uma observação.

**Distribuição Bidimensional** - é a representação de uma variável bidimensional  $(x_i, x_j)$ , com  $1 \leq i \leq n$  e  $x_j$  duas variáveis unidimensionais.

**Diagrama de Caule-e-Folhas** - o mesmo que Separador de Frequências.

**Diagrama de Extremos e Quartis** - é um diagrama que representa os valores extremos e os quartis de uma variável estatística.

**Distribuição de Frequências** - o mesmo que Tabela de Frequências.

**Estatística** - é o método que ensina a recolher, classificar, apresentar e interpretar um conjunto de dados numéricos.

**Estatística Descritiva** - ramo da Estatística que tem por finalidade descrever certas propriedades relativas a um conjunto de dados.

**Estatística Indutiva** - ramo da Estatística que procura inferir propriedades da população a partir de propriedades verificadas numa amostra da mesma.

**Fenômenos Independentes** - são fenômenos respeitantes à mesma variável que não têm qualquer ligação um com o outro.

**Frequência Absoluta (  $f_i$  )** - é o número de vezes que o valor de determinada variável é observado.

**Frequência Absoluta Acumulada (  $F_i$  )** - é a soma das frequências absolutas anteriores com a frequência absoluta deste valor.

**Frequência Relativa (  $f_{ri}$  )** - é o quociente entre a frequência absoluta do valor da variável e o número total de observações.

**Frequência Relativa Acumulada (  $F_{ri}$  )** - é a soma das frequências relativas anteriores com a frequência relativa desse valor.

**Função Cumulativa** - função que indica para cada valor real  $x$  a frequência absoluta (ou relativa) de observações com intensidade menor ou igual a  $x$ . A representação gráfica desta função é em forma de escada.

**Gráfico Circular** - é representado por um círculo que está dividido em sectores cujas amplitudes são proporcionais à frequência que lhe corresponde.

**Gráfico de Barras** - é constituído por barras, horizontais ou verticais, de comprimento proporcional à frequência.

**Histograma** - é um gráfico de barras em que a área destas é proporcional à frequência, não havendo espaço entre as mesmas. Só se utiliza em variáveis quantitativas contínuas.

**Média Aritmética Simples** - é o quociente da soma de todos os dados não classificados pelo número desses dados.

**Média Aritmética Ponderada** - é o quociente entre o somatório do produto de cada dado classificado pela sua frequência absoluta e o número desses dados.

**Mediana** - é o valor da variável, para dados não classificados, que ocupa a posição central da distribuição.

**Medidas de Dispersão** - é um conjunto de medidas (Amplitude, Variância e Desvio Padrão) utilizadas no estudo da variabilidade de uma determinada distribuição, permitindo obter uma informação mais completa acerca da "forma" da mesma.

**Medidas de Localização** - é um conjunto de medidas (Média, Mediana, Moda e Quartis) que representam de uma forma global um conjunto de dados.

Medidas de Tendência Central - o mesmo que Medidas de Localização.

**Moda ( m )** - observação que ocorre com maior frequência numa amostra.

**Nuvem de Pontos** - o mesmo que Diagrama de Dispersão.

**Organização dos Dados** - consiste em "resumir" os dados através da sua contagem e agrupamento.

**Pictogramas** - são gráficos onde se utilizam figuras ou símbolos alusivos ao problema em estudo.

**Planificação do Problema** - consiste na determinação de um processo para resolver o problema e, em especial, como obter informações sobre a variável em estudo.

**Polígono de Frequências** - são gráficos com aspecto de linhas quebradas. Constroem-se unindo por segmentos de reta os pontos médios das bases superiores dos retângulos de um histograma.

**População** - é um conjunto de seres com uma dada característica em comum e com interesse para o estudo.

**Quartis ( Q1 e Q3 )** - são os valores que dividem a distribuição em quatro partes iguais.

**Recenseamento** - o mesmo que Censo.

**Recolha de Dados** - é a primeira etapa depois de definido o problema em estudo.

**Reta de Regressão** - é a reta traçada sobre uma dada nuvem de pontos, sendo um modelo matemático que pretende descrever a relação existente entre duas variáveis unidimensionais de uma distribuição bidimensional.

**Relações Estatísticas** - são relações que se podem estabelecer entre determinadas variáveis de um problema em estudo.

**Separador de Frequências** - é um tipo de tabela que permite ter uma percepção imediata do aspecto global dos dados sem perda da informação contida na coleção dos dados inicial.

**Somatório (  $\Sigma$  )** - representa, de forma abreviada, uma soma.

**Sondagem** - é o estudo estatístico que se baseia numa parte da população, isto é, numa amostra que deve ser representativa dessa população.

**Tabela de Frequências** - são tabelas onde se apresentam os dados por classes e as frequências respectivas.

**Tamanho da Amostra** - é o número de elementos que constituem uma dada amostra.

**Unidade Estatística ou Indivíduo** - é cada um dos elementos da população.

**Variância** - é a medida que permite avaliar o grau de dispersão dos valores da variável em relação à média.

**Variáveis Contínuas** - são as variáveis que podem tomar qualquer valor de um determinado intervalo.

**Variáveis Discretas** - são as variáveis que podem tomar um número finito ou uma infinidade numerável de valores.

**Variáveis Qualitativas** - o mesmo que Atributos Qualitativos.

**Variáveis Quantitativas** - o mesmo que Atributos Quantitativos.

**ANEXO C - Mensagens retiradas de ATV\_01\_07 – AT2**

ALUNO	MENSAGEM
48	boa noite
50	boa
47	oi nene
48	Retrospectivo: baseado em dados de períodos passados.
47	vcs ja baixaram os lvros?
48	ainda nao
50	já
50	<a href="https://pt.slideshare.net/FClinico/tipos-de-estudos-epidemiolgicos-26672507">https://pt.slideshare.net/FClinico/tipos-de-estudos-epidemiolgicos-26672507</a>
56	Cara, é so pra mim que essas frase sao confusas? oO
50	pra mim também está, o professor não falou sobre isso ainda
47	só
48	a B acho que esta certa
50	a "C" também
48	a A ta confusa
56	é "B" pra mim ta certa tbm
48	a D tambem por causa do final
48	e outro sem esse fator.
47	cra
47	nao pode isso
47	aqui ta rodano norma
50	B, C
48	a A pode estar certa tambem
48	nao, estou enganado
48	e confuso
48	B,C
50	a D também
48	e esse outro sem esse fator
48	eu nao entendi cara
50	B, C e D
48	entendi agora
56	a C ta certa msm? onde vc achou
48	todo mundo condorda ?
50	a C mandei antes
50	achei no linkedin slideshare
50	O que acham?
56	vamo
50	viu a imagem?
50	concorda?
48	concordo
50	a A não entendi
56	ja respondi
56	A tbm nao deu boa
48	ela auxilia entrevista



48	mas sobre exames de registros eu nao achei nada
50	também não, apareceu algo sobre enfermagem
56	4min reponde ai time
50	a C deu uma bugada agora kkk
47	o que vcs acaham qu é?
48	eu ja respondi
47	o que é pra responder
50	A C eles dão o exemplo de uma doença que está relacionada à outra
56	então cara, isso que eu to achando mto estranho
56	msm pesquisando, nao tem nada muito relacionado com as alternativas
50	sim
50	tiramos a C?
48	sim
50	B e D?
50	Fechou?
48	melhor acertar parcialmente do que errar tudo
50	Foi mal ae time
50	confundi nós
50	Responderam?
47	vai renan
47	responde vado
47	viad
47	o
56	vcs
56	tem que responder
56	d novo
50	já respondi
56	eu tbm
50	que boxta
47	aeee
56	nice
48	nice
50	aaaaaeeee
50	ppp
50	que legal
50	agora aquele excel
56	A e E?
47	modal 17?
47	nosa
56	modal 13 n?
47	esuce
47	to viajendo
47	ta certo
47	respondi

50	sim
50	fechou
50	A e E

**APÊNCIDE A - Mensagens significativas retiradas de ATV\_01\_07 – AT3**

Aluno	Mensagem	Classificada como
47	Qual a fórmula?	temOpinioao, temTermoMatematico
47	Qual o valor inicial da variância?	temTermoMatematico
50	O desvio padrão é de quatro	temTermoMatematico, temAfirmacao
47	Se a variação está entre esses valores	temTermoMatematico
50	Aí vc tá falando de desvio padrão ou de variância?	temOpinioao, temTermoMatematico
47	Só alteraria de o valor fosse maior que 36 ou menor que 28	temOpinioao, temTermoMatematico
47	Variância	temTermoMatematico
56	pqp eu tava fazendo o calculo do desvio padrao	temTermoMatematico
47	Dado um conjunto de dados, a variância é uma medida de dispersão que mostra o quão distante cada valor desse conjunto está do valor central (médio).	temTermoMatematico
48	Mas então ta estranho mesmo, pq a distribuição de salários altera sim, se o mesmo valor for dividido por mais 2 funcionários é pra alterar	temOpinioao, temTermoMatematico
50	Como foi o cálculo que vcs fizeram?	temOpinioao, temTermoMatematico
47	Aquela fórmula que vc passou	temOpinioao, temTermoMatematico
56	16 - 100%	temTermoMatematico
56	15,2 - x	temTermoMatematico
56	x = 95%	temTermoMatematico
56	entao o novo resultado seria 95% ou seja 5% menor	temTermoMatematico, temOpinioao
48	A media não altera mas a variante do salario sim, precisaria aumentar proporcionalmente o salario	temTermoMatematico
48	Era vey pensa comigo, se aumentar o numero de funcionarios, não ia alterar a media de idade, mas altera o valor de distribuição de salario	temOpinioao, temTermoMatematico
47	A Á lógica era que o valor iria mudar	temTermoMatematico
48	38 = 100% -> logo aumento de 2 membros é = 4,8% de diferença de distribuição de salario kkk fiz de cabeça mas não façam isso em casa	temTermoMatematico, temOpinioao

47	Por ser um valor com com uma diferença pequena	temTermoMatematico
47	O 5% seria justo	temTermoMatematico, temOpinioao, temSugestao
50	To tentando calcular esse daqui agora	temTermoMatematico
56	a media é quanto?	temOpinioao, temTermoMatematico
50	12,54	temTermoMatematico
48	12,54	temTermoMatematico
47	Meu desvio padrão deu 2.7	temAfirmacao, temTermoMatematico
48	o meu deu 1,759	temAfirmacao, temTermoMatematico
47	A variação deu isso	temTermoMatematico, temAfirmacao
50	O meu deu 1,76	temAfirmacao, temTermoMatematico
56	o meu deu 1,76 tbm	temAfirmacao, temTermoMatematico
48	variancia deu 3,096	temAfirmacao, temTermoMatematico
50	3,09598	temTermoMatematico

**APÊNDICE B - Mensagens significativas retiradas de ATV\_01\_07 – AT2**

Aluno	Mensagem	Classificada como
48	boa noite	temSaudacao
47	oi nene	temSaudacao
48	Retrospectivo: baseado em dados de períodos passados.	temTermoMatematico
48	a B acho que esta certa	temExplicacao, temAfirmacao
48	e outro sem esse fator.	temTermoMatematico
48	a A pode estar certa tambem	temOpinio
56	a C ta certa msm? onde vc achou	temOpinio
48	todo mundo condorda ?	temOpinio
50	concorda?	temOpinio
48	Concordo	temConcordancia
48	mas sobre exames de registros eu nao achei nada	temTermoMatematico, temOpinio
47	o que vcs acaham qu é?	temOpinio
56	então cara, isso que eu to achando mto estranho	temOpinio, temExplicacao, temAfirmacao
56	msm pesquisando, nao tem nada muito relacionado com as alternativas	temTermoMatematico, temNegacao
50	tiramos a C?	temOpinio
50	Fechou?	temOpinio
48	melhor acertar parcialmente do que errar tudo	temOpinio, temSugestao
50	Responderam?	temOpinio
50	que legal	temEstimulo
56	A e E?	temOpinio
47	modal 17?	temTermoMatematico, temOpinio
56	modal 13 n?	temTermoMatematico, temOpinio
47	ta certo	temOpinio