



Ramon Bambini Moreira

# **Utilizando análise semântica para minerar implicações significantes em mapas conceituais**

Vitória, ES

2022

Ramon Bambini Moreira

## **Utilizando análise semântica para minerar implicações significantes em mapas conceituais**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Informática da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do Grau de Mestre em Informática.

Universidade Federal do Espírito Santo – UFES

Centro Tecnológico

Programa de Pós-Graduação em Informática

Orientador: Prof. Dr. Davidson Cury

Coorientador: Prof. M.Sc. Rodrigo Ruy Boguski

Vitória, ES

2022

Ficha catalográfica disponibilizada pelo Sistema Integrado de Bibliotecas - SIBI/UFES e elaborada pelo autor

---

B199u Bambini Moreira, Ramon, 1993-  
Utilizando análise semântica para minerar implicações significantes em mapas conceituais / Ramon Bambini Moreira. - 2022.  
82 f. : il.

Orientador: Davidson Cury.  
Coorientador: Rodrigo Ruy Boguski.  
Dissertação (Mestrado em Informática) - Universidade Federal do Espírito Santo, Centro Tecnológico.

1. Mapeamento conceitual. 2. Processamento de linguagem natural (Computação). 3. Mineração de dados (Computação). I. Cury, Davidson. II. Ruy Boguski, Rodrigo. III. Universidade Federal do Espírito Santo. Centro Tecnológico. IV. Título.

CDU: 004

---



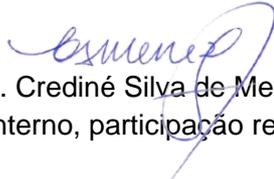
# ***Utilizando análise semântica para minerar implicações significantes em mapas conceituais***

***Ramon Bambini Moreira***

Dissertação submetida ao Programa de Pós-Graduação em Informática da Universidade Federal do Espírito Santo como requisito parcial para a obtenção do grau de Mestre em Informática.

Aprovada em 29 de julho de 2022.

Prof. Dr. Davidson Cury  
Orientador, participação remota



Prof. Dr. Crediné Silva de Menezes  
Membro Interno, participação remota

Prof. Dr. Alberto Nogueira de Castro Junior  
Membro Externo, participação remota

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO  
Vitória/ES, 29 de julho de 2022



UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO

**PROTOCOLO DE ASSINATURA**



O documento acima foi assinado digitalmente com senha eletrônica através do Protocolo Web, conforme Portaria UFES nº 1.269 de 30/08/2018, por  
DAVIDSON CURY - SIAPE 99992002  
Departamento de Informática - DI/CT  
Em 30/07/2022 às 10:38

Para verificar as assinaturas e visualizar o documento original acesse o link:  
<https://api.lepisma.ufes.br/arquivos-assinados/526972?tipoArquivo=O>



Documento assinado digitalmente

ALBERTO NOGUEIRA DE CASTRO JUNIOR

Data: 11/08/2022 22:48:27-0300

Verifique em <https://verificador.iti.br>

*A todos aqueles que de alguma forma estiveram e estão próximos de mim, fazendo esta vida valer cada vez mais a pena.*

## AGRADECIMENTOS

Em primeiro lugar agradeço a Deus por mais esta conquista. Por me dar ânimo, paciência, dedicação e força de vontade para enfrentar todas as dificuldades e situações adversas com as quais tive que lidar durante essa caminhada. Não foi nada fácil enfrentar tamanhos desafios em um momento de crise global como a pandemia da COVID-19. Desenvolver uma pesquisa de maneira remota, sem ter um contato diário com outros pesquisadores ou poder frequentar os laboratórios da universidade dificultou ainda mais todo esse desafio.

Sou imensamente grato à minha família, aos meus pais, Jamir e Kátia por terem me apoiado, dado muita força e carinho e me ajudado a chegar onde cheguei, e aos meus irmãos, Renan e Giseli. Agradeço também a minha noiva Luísa, por todo amor e incentivo nas horas difíceis.

Um obrigado mais que especial ao meu orientador, Dedê, e ao meu co-orientador, Rodrigo. Vocês me trouxeram lições e conhecimento que eu nunca imaginei adquirir, agradeço também pela paciência, disposição e boa vontade. Foi um prazer ter sido orientado por vocês. Assim também agradeço aos meus professores, que com muita competência e profissionalismo, me ensinaram o máximo que podiam da melhor forma possível, sou muito grato a vocês.

Por fim, dedico minha gratidão a todos colegas com quem frequentei salas de aula virtuais durante esses dois anos, vocês fizeram parte do aprendizado, contribuindo para que eu pudesse crescer como pessoa.

Muito obrigado!

*“Na vida, algumas vezes ganhamos e algumas vezes aprendemos.”*  
*(John C. Maxwell)*

## Resumo

Mapas conceituais são formas de representação gráfica que estabelecem relações entre conceitos. Eles podem ser utilizados na educação em diversas situações e finalidades: como recurso de aprendizagem, representação cognitiva, meio de avaliação, organização instrucional e compartilhamento de conhecimento. O conhecimento evidenciado em um mapa conceitual envolve implicações entre significações. Significações são tudo o que pode ser dito de um objeto, como uma descrição das suas propriedades, bem como tudo o que podemos observar nele. Além disso, uma significação ou implicação também refere-se a tudo o que podemos pensar dos objetos (classificá-los, estabelecer algum tipo de relacionamento, dentre outros).

Em sala de aula, os mapas conceituais podem ser aplicados como forma de mediar o ensino-aprendizagem, estimulando a aprendizagem significativa. Existem muitos desafios para que a utilização de mapas conceituais seja mais eficaz na medição da aprendizagem e avaliação dos processos cognitivos do aprendiz e suas interações com outros participantes da aprendizagem. A avaliação da aprendizagem é uma tarefa bem complexa, principalmente quando se almeja uma automatização do processo avaliativo, que demanda, dentre outras obrigações, uma formalização das estruturas de representação do conhecimento. Ademais, a avaliação de um mapa conceitual torna-se mais complexa quando o autor não representa o seu conhecimento, mas o conhecimento expresso em um texto ou em um mapa feito por outra pessoa.

Neste contexto, realizamos uma revisão da literatura entre os anos de 2015 e 2020 sobre o desenvolvimento de abordagens tecnológicas que auxiliam ou automatizam o processo de avaliação de mapas conceituais no ambiente educacional, através do uso de implicações significantes. Também buscamos identificar as limitações e reunir as melhores características dos trabalhos relacionados para propor nossa abordagem. Entre as limitações encontradas, constatamos que nenhuma das abordagens encontradas aplicava análise semântica na avaliação dos mapas conceituais.

A aplicação da análise semântica possibilita alcançar uma percepção mais profunda do suposto conhecimento representado nos mapas conceituais, buscando transformar indícios em evidências da aprendizagem, de fato, realizada pelo aprendiz, auxiliando o professor na avaliação dos mapas conceituais. Além do mais, uma atividade que envolva mapas conceituais em uma sala de aula resultaria na construção de diferentes mapas, e com isso, a avaliação se tornaria custosa e de grande esforço cognitivo para o professor. Por isso, a automatização desse processo é de grande valia.

Com o intuito de desenvolver uma arquitetura computacional capaz de efetuar análise semântica de mapas conceituais por intermédio das implicações significantes definidas por Piaget, esta pesquisa resultou em uma nova API, denominada MAPimp 2.0, que tem como objetivo de fornecer uma dimensão semântica mais profunda na análise das implicações significantes e de extrair informações mais precisas sobre a representação do conhecimento e a compreensão de um indivíduo a partir da representação de seu mapa. Essa nova versão traz a implementação de validações sintático-semânticas como novidade. A arquitetura é capaz de processar mapas conceituais escritos no idioma inglês e português. O desenvolvimento se deu usando técnicas de processamento de linguagem natural e modelos de redes neurais pré-treinados com bilhões de textos para a previsão de palavras e cálculo de similaridade semântica. A fim de obter uma análise quantitativa e qualitativa, a arquitetura conceitual foi aplicada em um ambiente de sala de aula, apresentando resultados satisfatórios.

**Palavras-chaves:** Mapas Conceituais. Implicações Significantes de Piaget. Processamento de Linguagem Natural. Análise Semântica.

## Abstract

Concept maps are forms of graphic representation that establish relationships between concepts. They can be used in education in different situations and purposes: as a learning resource, cognitive representation, means of evaluation, instructional organization and knowledge sharing. The knowledge evidenced in a conceptual map involves implications between meanings. Meanings are everything that can be said about an object, such as a description of its properties, as well as everything that we can observe in it. In addition, a meaning or implication also refers to everything we can think of objects (classifying them, establishing some kind of relationship, among others).

In the classroom, concept maps can be applied as a way of mediating teaching-learning, stimulating meaningful learning. There are many challenges for using concept maps to be more effective in measuring learning and evaluating the learner's cognitive processes and their interactions with other learning participants. The assessment of learning is a very complex task, especially when the objective is to automate the assessment process, which demands, among other obligations, a formalization of knowledge representation structures. Furthermore, the evaluation of a concept map becomes more complex when the author does not represent his knowledge, but the knowledge expressed in a text or in a map made by someone else.

In this context, we carried out a literature review between the years 2015 and 2020 on the development of technological approaches that help or automate the process of evaluating concept maps in the educational environment, through the use of significant implications. We also sought to identify limitations and gather the best features of related works to propose our approach. Among the limitations found, we found that none of the approaches found applied semantic analysis in the evaluation of concept maps.

The application of semantic analysis makes it possible to achieve a deeper perception of the supposed knowledge represented in the concept maps, seeking to transform evidence into evidence of the learning, in fact, carried out by the learner, helping the teacher in the evaluation of the concept maps. Furthermore, an activity that involves concept maps in a classroom would result in the construction of different maps, and with that, the evaluation would become costly and of great cognitive effort for the teacher. Therefore, the automation of this process is of great value.

In order to develop a computational architecture capable of performing semantic analysis of concept maps through the significant implications defined by Piaget, this research resulted in a new API, called MAPimp 2.0, which aims to provide a deeper semantic dimension in

the analysis. of significant implications and to extract more accurate information about the representation of knowledge and an individual's understanding from the representation of their map. This new version brings the implementation of syntactic-semantic validations as a novelty. The architecture is capable of processing concept maps written in English and Portuguese. The development took place using natural language processing techniques and pre-trained neural network models with billions of texts for word prediction and semantic similarity calculation. In order to obtain a quantitative and qualitative analysis, the conceptual architecture was applied in a classroom environment, presenting satisfactory results.

**Keywords:** Concept Maps. Piaget's Significant Implications. Natural Language Processing. Semantic Analysis.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Síntese do método de pesquisa. . . . .	22
Figura 2 – Síntese do processo de pesquisa. . . . .	24
Figura 3 – Exemplo de Mapa Conceitual. . . . .	28
Figura 4 – Os níveis de implicações significantes (DUTRA, 2018). . . . .	30
Figura 5 – Exemplo de Mapa conceitual sobre funcionamento do Motor a Explosão. . . . .	32
Figura 6 – Mapa conceitual representando as implicações significantes. . . . .	33
Figura 7 – Concentração de estudos por ano. . . . .	36
Figura 8 – Tabulação dos dados referente aos trabalhos selecionados. . . . .	37
Figura 9 – Modelo conceitual. . . . .	42
Figura 10 – Mapa de referência com avaliação das implicações significantes. . . . .	45
Figura 11 – Critérios de pontuação na avaliação do mapa conceitual. . . . .	51
Figura 12 – Mapa conceitual de um aluno avaliado pelas implicações significantes. . . . .	52
Figura 13 – Tela principal da API. . . . .	55
Figura 14 – Requisição do serviço com o método POST e o endereço da URL do serviço. . . . .	61
Figura 15 – Código JSON para ser requisitado ao serviço. . . . .	62
Figura 16 – Código JSON retornado pelo serviço com as implicações significantes e demais informações. . . . .	63
Figura 17 – Árvore de dependência para a frase do exemplo. . . . .	66
Figura 18 – Mapa de referência da atividade sobre Preservação Digital. . . . .	72

## LISTA DE TABELAS

Tabela 1 – <i>String</i> de busca. . . . .	35
Tabela 2 – Processo de seleção do estudo primário. . . . .	36
Tabela 3 – Meios de avaliação utilizados pelos estudos selecionados. . . . .	39
Tabela 4 – Técnicas de análise semântica empregadas nos estudos selecionados. . .	39
Tabela 5 – Análise dos Mapas para os experimentos EXP-BEGIN e EXP-END. . .	73
Tabela 6 – Análise sumarizada dos experimentos . . . . .	73

## LISTA DE ABREVIATURAS E SIGLAS

API	Application Programming Interface
BERT	Bidirecional Encoder Representations from Transformers
HTTP	Hypertext Transfer Protocol
JSON	JavaScript Object Notation
PLN	Processamento de Linguagem Natural
POS	Part of Speech Tagging
WWW	World Wide Web

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>17</b>
1.1	Contexto	17
1.2	Motivação	20
1.3	Hipóteses da Pesquisa	20
1.4	Questões da Pesquisa	21
1.5	Objetivo da Pesquisa	21
1.6	Método de Pesquisa	22
1.7	Processo de Pesquisa	23
1.8	Contribuição da Pesquisa	25
1.9	Produção Científica	26
1.10	Organização desta Dissertação	26
<b>2</b>	<b>MAPAS CONCEITUAIS E AS IMPLICAÇÕES SIGNIFICANTES</b>	<b>27</b>
2.1	Mapas Conceituais	27
2.2	Implicações Significantes	29
2.2.1	Categorias para a análise de Mapas Conceituais	30
2.3	Algumas considerações sobre o capítulo	33
<b>3</b>	<b>REVISÃO DA LITERATURA</b>	<b>34</b>
3.1	O Processo da Revisão	34
3.1.1	Questões de Pesquisa	34
3.1.2	Pesquisa sobre os Estudos Primários	35
3.2	Análises dos Resultados	37
3.3	Resultados e Discussão	37
3.4	Considerações sobre o capítulo	39
<b>4</b>	<b>O MODELO CONCEITUAL</b>	<b>41</b>
4.1	O Modelo	41
4.2	Análise semântica em mapas conceituais	44
4.2.1	Identificar Implicações locais	45
4.2.2	Identificar Implicações sistêmicas	46
4.2.3	Identificar Implicações estruturais	47
4.2.4	Análise semântica de implicações	47
4.3	Realçar Implicações Significantes	48

4.4	<b>Avaliação do Mapa Conceitual</b>	49
4.5	<b>Considerações sobre o capítulo</b>	52
5	<b>MAPIMP 2.0</b>	54
5.1	<b>Sobre o MAPimp 2.0</b>	54
5.1.1	Documentação da API	55
5.1.2	Exemplo de utilização	60
5.2	<b>SpaCy</b>	63
5.2.1	Tokenização	64
5.2.2	Lematização	64
5.2.3	<i>Part of Speech Tagging</i>	65
5.2.4	Vetores de Palavras e Similaridade Semântica	66
5.3	<b>BERT</b>	67
5.4	<b>Considerações sobre o capítulo</b>	69
6	<b>PROVA DE CONCEITO</b>	70
6.1	<b>Visão Geral</b>	70
6.2	<b>Experimento</b>	70
6.3	<b>Premissas do experimento</b>	71
6.4	<b>Verificação dos Resultados</b>	72
6.5	<b>Considerações sobre o capítulo</b>	74
7	<b>CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS</b>	76
7.1	<b>Considerações Finais</b>	76
7.2	<b>Trabalhos Futuros</b>	78
	<b>REFERÊNCIAS</b>	79

## CAPÍTULO 1

## INTRODUÇÃO

*Este capítulo apresenta uma visão geral da pesquisa realizada no decorrer deste trabalho, explicando ideias sobre contexto, motivação, hipóteses, questões, objetivos, métodos, processo, contribuições e produções. Essas explicações e discussões guiarão todos os capítulos subsequentes.*

## 1.1 Contexto

Desde o princípio das civilizações, o ser humano utiliza diversos recursos para simbolizar a realidade que o circunda. Ao produzir pinturas rupestres, o homem pré-histórico desenhava figuras que retratavam práticas do seu cotidiano. Ao trazer à análise esses artefatos arqueológicos, pode-se deduzir sobre o ambiente, a realidade, os detalhes da cultura e do modo de vida daqueles povos. A pintura, nesse caso, foi um elemento usado para a representação da realidade. A representação das coisas é atrelada ao conceito de substituição. Representando, criamos uma relação entre o que se apresenta e o signo, num ato de substituição, mas que não pode deixar de ser um gesto de conhecimento (LIMA; ALVARES, 2012).

Representar é a prática de empregar elementos simbólicos, como palavras, imagens, desenhos, esquemas, entre outros, para substituir um objeto, uma ideia ou um acontecimento. O ato de representar está correlacionado com o objeto, com a materialização e com o registro da simbologia que substitui o objeto ou ideia. Podem existir representações que estão desordenadas e nem por isso deixam de ser representações. Já a organização do conhecimento é o modo como ele é estruturado em assuntos onde se deseja a sua sistematização ordenada para atingir determinado propósito.

Na ciência da informação, quando se fala de organização e representação da informação e do conhecimento, a ênfase é dada aos processos de organização, utilizando-se de representações das informações e dos conhecimentos cujos elementos, para serem compreendidos, precisam de alguns conceitos preliminares (LIMA; ALVARES, 2012).

Existem diversas maneiras de representar a informação visualmente: gráficos, esquemas com chaves contendo itens e subitens, mapas mentais, mapas conceituais, fluxogramas, entre outros. O que diferencia mapas conceituais das demais formas de representação visual da informação é o uso de palavras/frases de ligação. Expressar a natureza da relação dos conceitos (se são “decorrentes”, “parte”, “consequência”, “exemplo”, “exceção”, etc.) leva o aluno a exteriorizar a compreensão que ele tem do conteúdo sendo trabalhado, exigindo maior concentração e esforço de sua parte, propiciando um desenvolvimento cognitivo

maior.

As vantagens de se utilizar mapas conceituais são: os conceitos mais inclusivos estão explícitos; os conceitos auxiliares e menos inclusivos estão inter-relacionados; o conhecimento é estruturado de maneira mais adequada a compreensão humana, considerando em posição de destaque os conceitos mais inclusivos. Já a sua desvantagem ocorre pela dificuldade de externalização e construção, visto que expõe a estrutura cognitiva do autor sobre o assunto. A clareza do autor sobre o tema fica evidente quando da sua construção. A sua construção sempre representa um desafio, dado que explicita (principalmente para si) a profundidade do conhecimento sobre o tema do mapa.

Assim, podemos afirmar que o mapa conceitual é uma ferramenta de representação gráfica não sequencial que permite fácil compreensão, construção e compartilhamento do conhecimento. Como se trata de uma técnica muito flexível, pode ser usada em diversas situações e diferentes finalidades: instrumento de análise do currículo, técnica didática, recurso de aprendizagem, meio de avaliação, dentre outros. Eles são utilizados como uma linguagem para descrição e comunicação de conceitos e seus relacionamentos, e foram originalmente desenvolvidos para dar suporte à aprendizagem significativa (AUSUBEL, 1978).

Segundo a teoria da aprendizagem significativa, os mapas conceituais são uma representação gráfica de um conjunto de conceitos construídos de tal forma que as relações entre eles são perceptíveis. Quando dois ou mais conceitos são relacionados por uma frase de ligação, é criada uma unidade semântica ou de significado, que chamamos de proposição.

Para Piaget e Garcia (1989), desde os níveis mais básicos de pensamento há implicações entre significações. Quando um aprendiz define uma relação entre dois conceitos durante a criação de seu mapa conceitual, ele realiza, em última análise, uma implicação significativa.

Mapas conceituais como ferramentas tecnológicas podem ser usados em sala de aula como forma de mediar o ensino-aprendizagem, estimulando a aprendizagem significativa. Existem inúmeros desafios para que a utilização de mapas conceituais seja mais eficaz na medição da aprendizagem e avaliação dos processos cognitivos do aprendiz e suas interações com outros participantes da aprendizagem.

A avaliação da aprendizagem é uma tarefa bem complexa, principalmente quando se almeja uma automatização do processo avaliativo, que demanda, dentre outras obrigações, uma formalização das estruturas de representação do conhecimento. Além disso, é desejado que um processo avaliativo seja utilizado como uma prática da construção do conhecimento, melhorando tanto o ensino quanto a aprendizagem (NOVAK; GOWIN; BOB, 1984). Além disso, a avaliação de um mapa conceitual torna-se mais complexa quando o autor não representa o seu conhecimento, mas o conhecimento expresso em um texto ou em um

mapa feito por outra pessoa.

Como instrumento de avaliação da aprendizagem, mapas conceituais podem ser usados para se obter uma visualização da organização conceitual que o aprendiz atribui a um dado conhecimento. Trata-se basicamente de uma técnica não tradicional de avaliação que busca informações sobre os significados e relações significativas entre conceitos-chave de um domínio segundo o ponto de vista do aluno. É mais apropriada para uma avaliação qualitativa, formativa, da aprendizagem (MOREIRA, 2012).

Diversos estudos no domínio da Informática na Educação têm investigado o problema de avaliar a aprendizagem de alunos a partir da comparação do mapa conceitual de cada aprendiz com um mapa definido como referência, através de ferramentas computacionais (BOSS; MATOS; ANDRADE, 2019).

Trabalhos desenvolvidos nessa linha de pesquisa (MOREIRA; AGUIAR; CURY, 2019), (RIOS, 2017), (RIOS; CURY; DUTRA, 2015) e (DUTRA, 2018), têm revelado, com passar dos anos, que automatizar o processo da avaliação da aprendizagem por meio da análise de mapas não é um procedimento simples e trivial. Muito se deve a certas questões que são levadas em consideração quando é feita a comparação entre mapas conceituais, como é o caso da análise sintática, para reconhecimento de padrões estruturais da linguagem em que o mapa foi construído e da análise semântica, para o reconhecimento dos significados das sentenças obtidas.

A análise semântica, como sempre, objetiva uma percepção mais profunda do suposto conhecimento representado no mapa, buscando transformar indícios em evidências da aprendizagem, de fato, realizada pelo aprendiz, auxiliando o professor na avaliação dos mapas conceituais.

Nesse contexto, notamos o desenvolvimento de abordagens tecnológicas que auxiliam ou automatizam o processo de avaliação de mapas conceituais no ambiente educacional, através do uso de Implicações Significantes. Essas abordagens adotam diferentes técnicas de Processamento de Linguagem Natural (PLN). Porém, podemos considerar que os seus resultados ainda não são satisfatórios e apresentam algumas limitações que serão discutidas a seguir.

Esta pesquisa é uma continuação de estudos iniciados anteriormente por Moreira (2019), Rios (2017), Dutra (2018) e Dutra et al. (2006). No entanto, essas abordagens não aplicam análise semântica na avaliação de mapas conceituais. Tais pesquisas fundamentam a avaliação apenas nas análises sintáticas e/ou topológicas do mapa conceitual, sem considerar a significação ou sentidos das relações expressas entre conceitos.

No contexto da pesquisa, adotamos o uso de mapas conceituais de estilo científico, que são regidos por duas regras básicas: o mapa pode conter apenas conceitos, e sempre há um verbo na relação entre os conceitos. Um mapa conceitual de estilo científico é

direcionado para uma finalidade específica, como avaliação e suporte para aprendizagem, representação e sumarização de texto, entre outros (AGUIAR; CURY; ZOUAQ, 2018).

## 1.2 Motivação

Os mapas conceituais podem ser utilizados na educação em diversas situações e finalidades: como recurso de aprendizagem, representação cognitiva, meio de avaliação, organização instrucional e compartilhamento de conhecimento.

Destarte, os mapas podem ser utilizados como instrumento de avaliação da aprendizagem, uma vez que os professores os utilizam para verificar o nível de compreensão do aluno, para analisar o conhecimento médio de uma turma, para identificar conceitos e significados mal assimilados ou explicitados e compartilhados sobre um domínio de estudo.

No entanto, um mapa conceitual representa o conhecimento do próprio autor. Assim, uma atividade que envolva mapas conceituais em uma sala de aula resultaria na construção de diferentes mapas, e com isso, a avaliação se torna custosa e de grande esforço cognitivo para o professor.

No ano de 2018, uma primeira abordagem foi desenvolvida por [Moreira \(2019\)](#) para a análise das implicações significantes de Piaget. Como na época tratava-se de um projeto iniciativo, as análises sintáticas para a identificação de verbos eram feitas com apoio de bases de verbos locais e verbos transitivos diretos. Desse modo, não considerávamos o contexto da frase em que o verbo estava inserido para poder definir sua transitividade. Além disso, a abordagem carecia de uma avaliação semântica das proposições para verificação das relações estabelecidas pelo aluno. Eram contabilizadas apenas as quantidades de cada tipo de implicação que o aluno formulou em seu mapa, sem considerar se estavam corretas ou próximas semanticamente daquilo que o professor esperava, quando comparado com o mapa de referência. Apesar disso, a abordagem inicial alcançou resultados promissores: 100% na identificação das implicações locais e estruturais e 90% nas implicações sistêmicas, conforme a análise do professor.

Deste modo, a principal motivação que conduziu o desenvolvimento desta pesquisa foi o de aprimorar a primeira abordagem, aplicando de análise sintática e semântica para avaliar mapas conceituais, visando proporcionar uma avaliação mais apurada do conhecimento representado pelo aluno e evidenciar possíveis insuficiências na aprendizagem, sendo uma proposta de um novo modo de avaliação do processo de aprendizagem.

## 1.3 Hipóteses da Pesquisa

Com base nos problemas apresentados na Seção 1.1, os seguintes argumentos foram formulados como hipóteses:

- (i) É possível criar uma ferramenta capaz de identificar implicações significantes, utilizando análise semântica, em mapas conceituais de estilo científico de forma a aprimorar a avaliação da aprendizagem.
- (ii) A utilização de modelos de linguagem pré-treinados para processamento de linguagem natural conseguem dar uma dimensão semântica na avaliação de mapas.
- (iii) A comparação de similaridade de palavras entre os mapas dos alunos de uma mesma turma podem evidenciar um fechamento contextual, de forma a ampliar a dimensão semântica.

## 1.4 Questões da Pesquisa

A seguir, formulamos as principais questões que esta pesquisa pretende investigar:

- (i) Existe alguma ferramenta ou serviço da web disponível para a avaliação semântica automática de mapas conceituais?
- (ii) Quais são as melhores técnicas e métodos atuais para fazer a verificação semântica de mapas conceituais?
- (iii) Como avaliar a aprendizagem dentro de um processo de construção de mapas conceituais considerando as suas semânticas?
- (iv) É possível medir o aprendizado do aluno a partir do uso das implicações significantes?
- (v) É possível identificar a evolução da aprendizagem do aluno ao longo do curso de uma disciplina utilizando implicações significantes?
- (vi) É possível evidenciar conceitos e relações não percebidos por um aluno os quais o professor esperava?

## 1.5 Objetivo da Pesquisa

O objetivo geral desta pesquisa é desenvolver uma arquitetura computacional para identificar as implicações significantes de Piaget aplicando análise semântica. Em outras palavras, extrair informações semânticas sobre a representação do conhecimento em mapas conceituais a fim de buscar compreensão acerca do que o aluno representou ou o que ele sabia no momento em que produziu seu mapa, a partir da análise significativa das proposições construídas por ele. A proposta é suportada por várias técnicas distintas de Processamento de Linguagem Natural (PLN).

Além disso, como objetivos secundários, a pesquisa ainda será capaz de: auxiliar

o aluno na construção sintática do mapa conceitual de forma correta (AGUIAR; CURY, 2016), checando eventuais inconsistências que podem estar presentes, sugerindo correções e validando construções sintáticas e semânticas. Tudo isto permitirá identificar perfis de alunos em uma turma e realizar uma medida do aprendizado de cada estudante em relação a um assunto, além de, poder agrupar alunos que pensam de maneira semelhante.

Para tanto, serão considerados os seguintes objetivos específicos:

- Identificar relações semânticas nos mapas conceituais com a utilização de modelos treinados por redes neurais;
- Identificar relações semânticas de similaridade com palavras utilizadas pelo professor no mapa de referência;
- Identificar relações semânticas de similaridade com palavras utilizadas pelo restante da turma;
- Definir métricas para medir o aprendizado dos alunos;
- Definir uma arquitetura conceitual que poderá ser utilizada como ferramenta ou serviço da web para análise semântica de implicações significantes.

## 1.6 Método de Pesquisa



Figura 1 – Síntese do método de pesquisa.

Fonte: Elaborado pelo autor.

A **Natureza** da pesquisa é definida como **Aplicada**, seguindo a classificação proposta por Ander-Egg (1978), uma vez que esse estudo tem interesse prático e os resultados são aplicados na solução de problemas reais como uma ferramenta que analisa a semântica de mapas conceituais de maneira automática, tornando-se um instrumento de avaliação da aprendizagem.

A **Abordagem do Problema**, seguindo a classificação proposta por Sampieri, Collado e Lucio (2013), é definida como **Qualitativa** para a obtenção do resultado da

assistência à aprendizagem por meio da avaliação semântica e como **Quantitativa** para a análise objetiva e comparação de resultados.

O **Objetivo**, seguindo a classificação proposta por Gil (2008), pode ser classificado como **Exploratório**, uma vez que a literatura é revisada a fim de fornecer uma visão geral do assunto, permitindo um maior entendimento do pesquisador sobre o tema pesquisado, uma vez que ainda há poucos estudos nessa área dentro do contexto de mapas conceituais.

O **Procedimento Técnico**, seguindo a classificação proposta por Gil (2008), é definido como **Bibliográfico**, **Estudo de Caso** e **Experimental**, uma vez que utilizamos referenciais teóricos para coletar informações, aplicar variáveis para observação de efeitos durante o desenvolvimento da abordagem, e estudar a influência da análise semântica em mapas conceituais no processo de avaliação da aprendizagem.

O **Método Científico**, conforme a classificação proposta por Marconi, Lakatos et al. (2007), é definido como **Indutivo** por considerarmos o conhecimento empírico e a experiência para extrair soluções. Assim, partimos de observações concretas sobre o processo e, em seguida, generalizamos a solução em conclusões prováveis.

Por último, o **Método de Pesquisa**, também seguindo a classificação proposta por Marconi, Lakatos et al. (2007), que propõe uma visão mais prática com uma finalidade restrita, é definido como **Tipológico**, pois determinamos as características de uma nova abordagem a partir da classificação e comparação de abordagens semelhantes, e **Estrutural**, com a investigação do fenômeno concreto na avaliação semântica de mapas conceituais, chegamos a um nível abstrato por meio da arquitetura conceitual, e voltamos ao concreto implementando a solução.

## 1.7 Processo de Pesquisa

A seguir, são exibidas as etapas do processo investigativo aplicado durante esta pesquisa. Para isso, seguimos como exemplo a proposta de Quivy e Campenhoudt (1992), de modo que o processo investigativo sucedeu-se em três fases e seis etapas. A Figura 2 sintetiza todo o processo e destaca os resultados produzidos em cada etapa, que são consideradas como entrada para as etapas subsequentes.

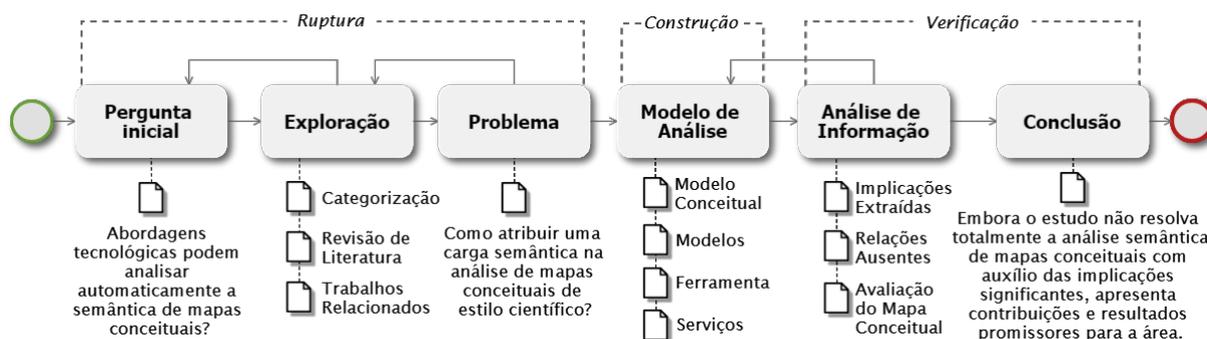


Figura 2 – Síntese do processo de pesquisa.

Fonte: Elaborado pelo autor.

A fase da **Ruptura** quebra os preconceitos e as falsas evidências. Esta fase resultou em uma revisão da literatura sobre as abordagens tecnológicas para análise semântica de mapas conceituais e levou à criação de uma categorização representando os vários métodos e características para a avaliação da significação nos mapas conceituais. Nesta fase, definimos as seguintes etapas:

- (i) **Pergunta Inicial:** Abordagens tecnológicas podem analisar automaticamente a semântica de mapas conceituais?
- (ii) **Exploração:** A exploração do assunto foi realizada pelos seguintes métodos:
  - Revisão da literatura (Seção 3.1) conduzida por um mapeamento sistemático na Biblioteca Digital IEEEExplore, Biblioteca Digital ACM e Elsevier ScienceDirect. Foram coletadas 223 publicações, das quais 63 foram pré-selecionadas, das quais 19 apresentaram abordagem relevante para nosso estudo.
  - Investigação de recursos tecnológicos disponíveis para realização de tarefas de PLN (APIs, serviços, etc..).
  - Observação e análise da informação catalogada, a partir da categorização criada nas abordagens tecnológicas, para a análise semântica de mapas conceituais.
- (iii) **Problema:** O problema foi estabelecido com base nos resultados obtidos na etapa de Exploração. Aplicamos a categorização em abordagens semelhantes para identificar as características de cada abordagem e seus pontos positivos e negativos. Portanto, definimos o problema “Como atribuir uma carga semântica na análise de mapas conceituais de estilo científico?”. Além disso, ainda era preciso adaptar as técnicas para o contexto das implicações significantes definidas por Piaget.

A fase de **Construção** demonstra a lógica do caso estudado e constrói suas proposições, plano de pesquisa e operações. Este ato resultou em um modelo conceitual, modelos para as línguas inglesa e portuguesa, e um serviço para a análise semântica de

mapas conceituais utilizando implicações significantes. Nesta fase, definimos a seguinte etapa:

- (i) **Modelo de Análise:** Desenvolvido a partir de observações e experimentos transformados de forma sistemática. Com base nas informações coletadas na fase de Ruptura, um modelo conceitual foi definido e os recursos necessários para a implementação do modelo foram desenvolvidos e adaptados, resultando em uma arquitetura orientada a serviços. Por fim, foi implementada uma ferramenta seguindo a arquitetura proposta.

A fase de **Verificação** analisa as informações obtidas pelo modelo criado na etapa de Construção. Essa fase resultou na verificação dos resultados apresentados pela ferramenta por meio de precisão, comparação e questionário. Nesta fase, definimos as seguintes etapas:

- (i) **Análise da Informação:** Verifica se os resultados observados correspondem aos esperados. Para tanto, utilizamos a Análise Estatística para comparar as implicações significantes identificadas pelo modelo criado com aquelas assinaladas por peritos humanos; Verificação das relações que estão ausentes no mapa do aluno, mas que eram esperadas pelo professor ao elaborar a atividade; Comparação entre a nota atribuída ao mapa conceitual pela ferramenta com a nota atribuída pelo professor.
- (ii) **Conclusão:** Com base nas informações e análises realizadas durante a pesquisa, concluímos que a avaliação semântica de um mapa conceitual é uma tarefa difícil até mesmo para especialistas humanos, e embora o estudo não tenha resolvido totalmente essa questão, apresenta contribuições e resultados promissores para a área.

## 1.8 Contribuição da Pesquisa

As principais contribuições que este estudo traz para as comunidades acadêmica e científica são as seguintes:

- (i) Detecção de inconsistências sintáticas no mapa conceitual;
- (ii) Método para identificar relações semânticas não observáveis em mapas conceituais;
- (iii) Uma arquitetura conceitual para dar suporte a atividades com mapas conceituais.
- (iv) Uma metodologia para a avaliação do aprendizado baseado em mapas conceituais;
- (v) Uma ferramenta que resulta em uma economia cognitiva do professor na análise de mapas construídos por seus estudantes, de forma a facilitar uma aprendizagem adaptada ao perfil do estudante;

- (vi) API contendo bibliotecas para análise semântica nos idiomas inglês e português, além de serviço publicamente disponível para consulta e expansão;
- (vii) Uma extensão disponibilizada para um ambiente de aprendizagem assistida por tutores inteligentes sobre mapas conceituais (BOGUSKI; CURY; GAVA, 2019), de modo que a identificação semântica de implicações fornece indícios de aprendizado ao tutor para que possa fazer intervenções com o aluno.

## 1.9 Produção Científica

Moreira, Ramon B.; Boguski, Rodrigo R.; Cury, Davidson (2021). **Utilizando análise semântica para descobrir implicações significantes em mapas conceituais**. Simpósio Brasileiro de Informática na Educação, 32. , 2021, Online. Porto Alegre: Sociedade Brasileira de Computação, 2021. p. 123-134. DOI: <https://doi.org/10.5753/sbie.2021.218489>.

## 1.10 Organização desta Dissertação

A pesquisa desenvolvida no decorrer desta dissertação está dividida em X capítulos. Os capítulos que seguem esta Introdução são:

**Capítulo 2:** Explora o contexto dos mapas conceituais e o propósito das implicações significantes de Piaget, que são as questões-chave para a compreensão desta pesquisa.

**Capítulo 3:** Este capítulo tem como objetivo apresentar um mapeamento sistemático das abordagens tecnológicas para a análise semântica de mapas conceituais.

**Capítulo 4:** Este capítulo apresenta a proposta de um modelo conceitual que aplica-se em mapas conceituais escritos em inglês ou português, abordando técnicas, ferramentas, arquitetura e estratégias utilizadas.

**Capítulo 5:** Este capítulo apresenta o MAPimp 2.0, uma nova versão da API Python que fornece um conjunto de bibliotecas de processamento de linguagem natural. Esta API é um dos resultados da pesquisa.

**Capítulo 6:** Apresenta o experimento realizado para a validação do modelo conceitual apresentado no Capítulo 4, apresentando o contexto, condições e pré-requisitos necessários.

**Capítulo 7:** Apresenta as considerações finais e discute trabalhos futuros.

## CAPÍTULO 2

## MAPAS CONCEITUAIS E AS IMPLICAÇÕES SIGNIFICANTES

*Neste capítulo, exploramos o contexto dos mapas conceituais e o propósito das implicações significantes de Piaget, que são as questões-chave para a compreensão desta pesquisa. Este capítulo está organizado da seguinte forma: a Seção...*

## 2.1 Mapas Conceituais

O mapa conceitual é uma forma de representação gráfica que estabelece relações entre conceitos. Trata-se de técnica que foi desenvolvida por Joseph Novak e seus colaboradores na Universidade de Cornell, nos Estados Unidos, em meados da década de 1970. Novak (NOVAK; CANAS, 2006) a define como uma ferramenta para organizar e representar o conhecimento, já que a estrutura cognitiva de um sujeito pode ser assimilada como um conjunto de conceitos que se relacionam entre si, a fim de formar proposições significativas.

De um modo geral, os mapas conceituais são compostos por conceitos, que aparecem em caixas, e relações entre conceitos, que são designadas por linhas rotuladas que os interligam por meio de setas direcionais. As palavras sobre as linhas determinam os relacionamentos entre dois conceitos. Para esta pesquisa, o nosso interesse está em mapas conceituais de estilo científico: conceitos são definidos por substantivos, e frases de ligação devem ter um verbo ou uma locução verbal (AGUIAR; CURY; ZOUAQ, 2018). Dois ou mais conceitos ligados por uma relação, criam uma unidade semântica, que chamamos de proposição.

Toda proposição estabelece uma verdade, um fato, destacável e compreensível por si só. As proposições são uma característica particular dos mapas conceituais em comparação com outras estruturas similares, tais como mapas mentais ou fluxogramas. De acordo com Novak, um mapa conceitual é uma estrutura arborescente hierárquica, onde os conceitos mais gerais ou inclusivos aparecem no topo e os mais específicos, nas partes mais baixas da árvore. Novo conhecimento é ancorado ao velho por afinidade semântica (RIOS; CURY; DUTRA, 2015).

Com base na teoria da aprendizagem significativa de Ausubel (1978), o mapa conceitual é uma representação gráfica de um conjunto de conceitos construídos de modo que as relações entre eles são evidentes. A aprendizagem é dita significativa quando uma nova informação (conceito, ideia, proposição) adquire novos significados para o indivíduo quando ela se ancora em aspectos relevantes da estrutura cognitiva preexistente

do indivíduo, isto é, para que a ancoragem de novos conceitos seja significativa, a estrutura cognitiva do indivíduo deve ter os conceitos pré-existentes necessários.

Como os recursos de aprendizagem são resultado de parcerias entre estudantes e professores, percebe-se que mapas conceituais também podem auxiliar como uma importante referência para os alunos, advertindo-os para as inúmeras possibilidades de se aprimorar qualitativamente os conceitos e relações cognitivas nos seus processos de construção do conhecimento.

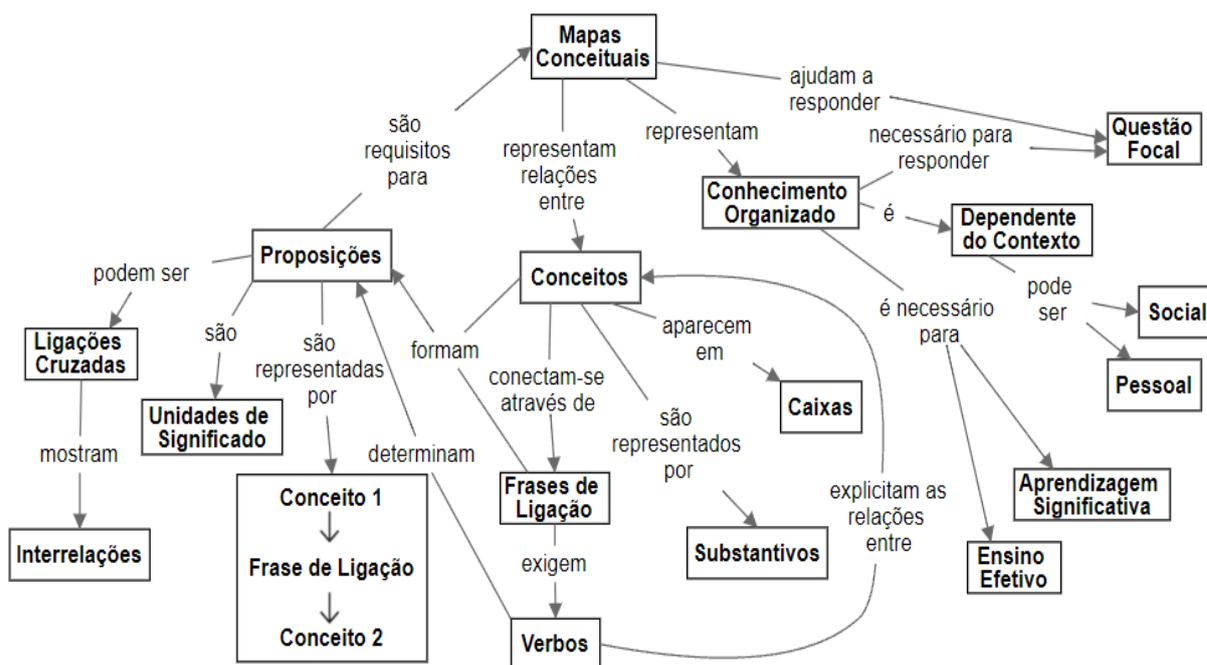


Figura 3 – Exemplo de Mapa Conceitual.

Fonte: Elaborado pelo autor.

A Figura 3 sintetiza todo o conteúdo abordado nos parágrafos anteriores e indica os elementos básicos necessários em um mapa conceitual. Observando a figura, podemos constatar que a organização hierárquica dos conceitos é estabelecida pelo posicionamento deles no mapa. Geralmente, conceitos de abrangência mais genérica localizam-se na parte superior do mapa, ao passo que aqueles mais específicos aparecem na parte inferior. Já as setas indicam a sequência e o direcionamento da construção do conhecimento.

O mapa conceitual também é elaborado a partir de uma única questão focal, que funda o conhecimento relevante para responder a uma questão de forma a contextualizar o mapa, embora possa abranger diferentes domínios ou segmentos. As ligações cruzadas estabelecem relações explícitas entre conceitos de domínios diferentes ou distantes.

## 2.2 Implicações Significantes

A aprendizagem pode ocorrer de diversas formas, por exemplo, ao olhar, tocar, sentir ou ouvir um objeto aprendemos alguma coisa sobre ele. Ao nos depararmos com algo novo, buscamos saber de suas propriedades, desejamos conhecer e entender suas funções e operações, assim como do que ele é capaz de fazer. Posteriormente, procuramos entender como essas operações ocorrem ou como aplicá-las de maneira diferente, e investigar do que ele é composto. Por último, descobrimos possíveis interações capazes de serem feitas com outros objetos. E assim, à medida que o conhecimento se forma, os diferentes níveis de complexidade do conhecimento vão se criando.

Piaget e Garcia (1989) classificaram esses níveis de conhecimento e ao conjunto deles denominaram implicações significantes. Elas são capazes de nos permitir identificar o grau de conhecimento de um indivíduo, sobre um dado domínio em um determinado momento.

O mapa conceitual da Figura 4 mostra, segundo o modelo piagetiano, uma distinção evolutiva dos níveis de implicação significativa.

Esses níveis resultam do desenvolvimento de estruturas que permitem a um indivíduo, a partir de um longo processo de construção de relações lógicas elaborado em função das situações em que se vê confrontado, expandir sua compreensão sobre os objetos do conhecimento. Desta forma, Piaget afirma que há uma Lógica das Significações que precede a Lógica Formal e que desde a ação da criança até suas operações em pensamento há uma correspondência entre as operações formais e as implicações significantes. O que ocorre é que as operações formais estão suficientemente diferenciadas para permitir a sua combinação em um sistema de conjunto. No caso das implicações significantes, é possível observar as mesmas operações formais em ações realizadas em contextos muito particulares sem, contudo, ser necessariamente generalizáveis ou compatibilizadas a outros contextos (DUTRA, 2018).

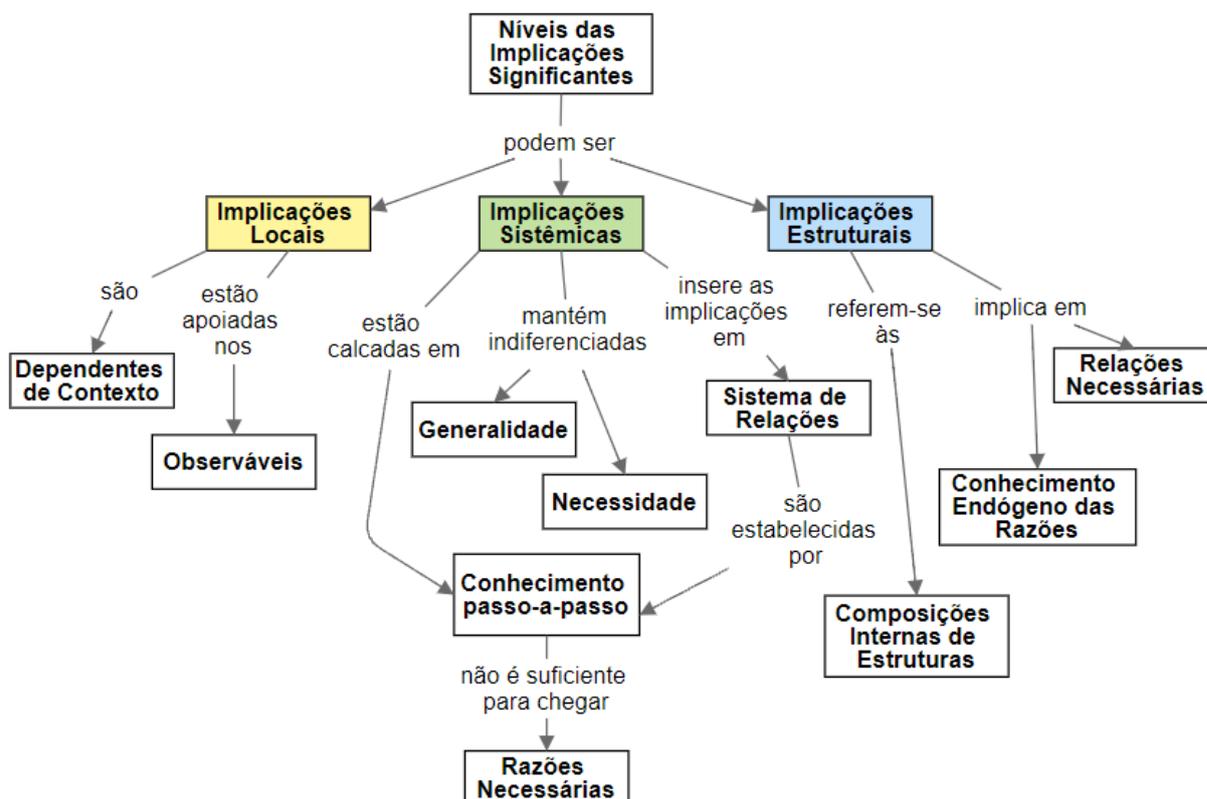


Figura 4 – Os níveis de implicações significantes (DUTRA, 2018).

Uma sentença tipo implicação transitiva  $A \rightarrow B \rightarrow C$ , na abordagem de Piaget, significa que se pelo menos uma significação de  $C$  está contida em qualquer uma de  $B$  está, por conseguinte, também contida em qualquer de  $A$ . Significação é tudo o que pode ser dito de um objeto, como uma descrição das suas propriedades, bem como tudo o que podemos observar nele. Além disso, uma implicação também é tudo o que podemos pensar dos objetos (classificá-los, estabelecer algum tipo de relacionamento etc.) (RIOS; CURY; DUTRA, 2015).

Quando se define uma relação entre dois conceitos (expressa por uma frase de ligação) durante a construção de um mapa conceitual, ao mesmo está se formando, em última análise, uma implicação significativa. Com um conjunto de implicações, podemos identificar o grau de conhecimento de um indivíduo em um determinado domínio.

### 2.2.1 Categorias para a análise de Mapas Conceituais

Nos itens subsequentes será apresentada uma adaptação (DUTRA; FAGUNDES; CAÑAS, 2004) da teoria das implicações significantes com o objetivo de analisarmos os mapas conceituais, com especial destaque para as frases de ligação.

- **Implicação Local** – pode ser definida quando o conhecimento não ultrapassa as propriedades observáveis do objeto em um determinado contexto. Uma implicação

local somente caracteriza um objeto. Elas podem qualificar um conceito ou estabelecer relações de parte do todo. Em um mapa conceitual, as implicações locais geralmente aparecem nas proposições com frases de ligação que usam verbos tais como é, tem, contém, possui etc., ou em predicados que qualificam o conceito.

- **Implicação Sistêmica** – generalizações e propriedades não diretamente observáveis começam a aparecer neste nível. A implicação sistêmica evidencia características do objeto que são notadas ou resultantes de alguma ação sobre ele. Esta implicação manifesta as relações de causa e efeito sem, contudo, revelar o por que isso ocorre. Estas relações ocorrem, em geral, pelo estabelecimento de uma relação do objeto em estudo com algum outro objeto. Para essa implicação, geralmente a causa e o efeito são, respectivamente, o sujeito e o objeto direto do verbo ou com o objeto indireto. São exemplos de verbos nocionais, que exercem ação ou evocam relação causal tais como causar, criar, fazer, gerar, disparar, produzir, emitir, resultar e assim por diante.
- **Implicação Estrutural** – amplia a implicação sistêmica, pois contém as razões e explicações da relação causa-efeito. Piaget fala da compreensão das razões endógenas e da descoberta das relações necessárias para o efeito ocorrer. Assim, mais do que o conhecimento de causas e efeitos, as implicações estruturais estabelecem condições que são essenciais para certas declarações, distinguindo-as daquelas que são apenas suficientes. A soma das ações de vários objetos relacionados à ocorrência de mais uma ação, estabelece a existência de uma implicação desse novo nível (RIOS; CURY; DUTRA, 2015). Nos mapas, elas são identificadas a partir de relações semânticas válidas entre duas ou mais proposições que convergem para formação de um novo conceito.

Em [Moreira, Aguiar e Cury \(2019\)](#), desenvolvemos uma técnica capaz de identificar as implicações significantes sobre os mapas conceituais e representá-las em uma ferramenta de apoio à aprendizagem. As implicações são identificadas por meio de cores nas setas das proposições, para que possam ser visualizadas, sendo amarelo para implicações locais, verde para sistêmicas e azul para estruturais. Eventuais proposições que não forem classificadas por nenhuma das implicações aparecem por padrão na cor preto.

A [Figura 5](#) reproduz um mapa conceitual acerca do funcionamento do motor a combustão. Nesse exemplo, quando dizemos que o motor tem cilindro e que a combustão é uma reação química, estamos fazendo implicações locais, já que isso pode ser registrado do objeto apenas a partir da observação de seu contexto e de seus atributos. Quando dizemos que a energia química, liberada da reação química da combustão, é transformada em energia mecânica, que, por sua vez, faz as rodas do carro movimentarem-se, estamos denotando que são implicações sistêmicas. Para chegarmos a uma implicação estrutural

precisaríamos explicar, por exemplo, que para que ocorra a combustão que ocorre dentro do cilindro do veículo, é necessário a atuação conjunta de combustível, ar e faísca.

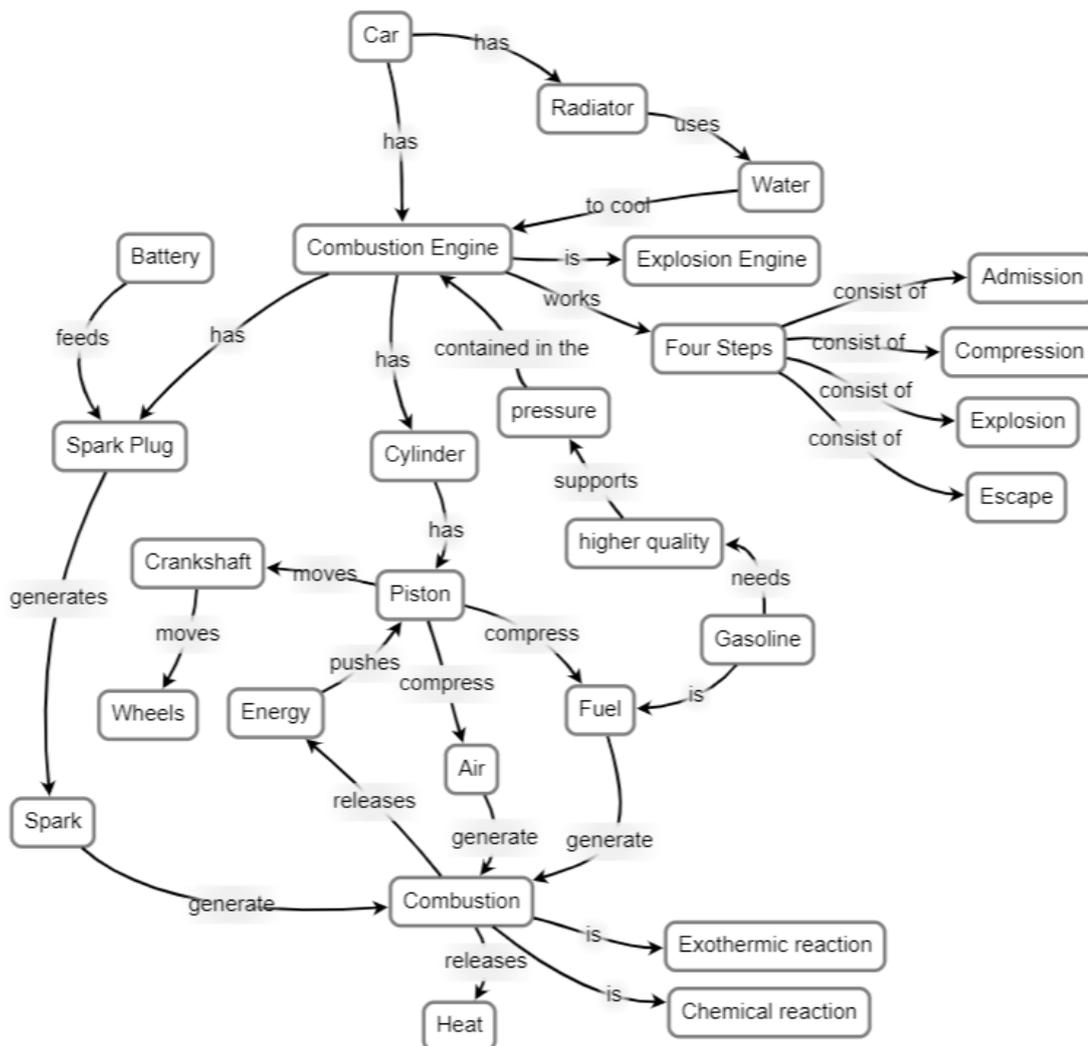


Figura 5 – Exemplo de Mapa conceitual sobre funcionamento do Motor a Explosão.  
 Fonte: Elaborado pelo autor.

A Figura 6 demonstra o mapa conceitual da Figura 5 representado com as implicações significantes de Piaget identificadas pelo arcabouço, em que as relações em amarelo, verde e azul evidenciam, respectivamente, implicações locais, sistêmicas e estruturais.

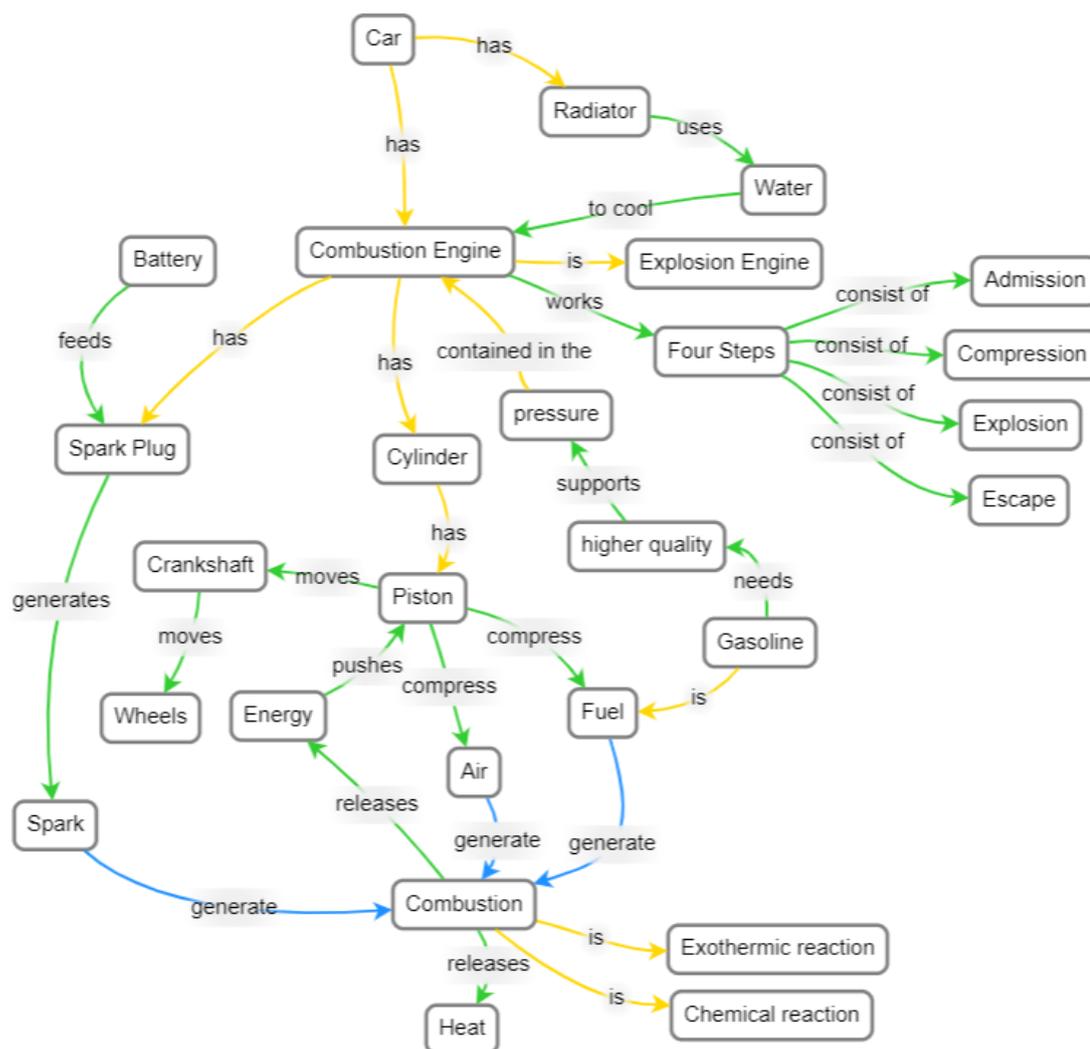


Figura 6 – Mapa conceitual representando as implicações significantes.

Fonte: Elaborado pelo autor.

## 2.3 Algumas considerações sobre o capítulo

Este capítulo trouxe conceitos-chave para a compreensão desta pesquisa com relação à análise das implicações significantes de Piaget, aplicadas ao contexto dos mapas conceituais.

Ressaltamos a importância do conhecimento das implicações significantes, que serão o cerne desta pesquisa para a realização da análise semântica, visto que o conhecimento envolve implicações entre significações, atribuídas às propriedades, aos objetos, aos conceitos e às próprias ações do sujeito. As significações resultam de suas aplicações, ou seja, das relações e das coordenações estabelecidas pelo sujeito. Além disso, uma análise aprofundada das implicações significantes pode nortear os próximos passos do ensino-aprendizagem.

O próximo capítulo apresenta uma revisão da literatura sobre a respeito da aplicação de análise semântica para a avaliação de mapas conceituais.

## CAPÍTULO 3

## REVISÃO DA LITERATURA

*Este capítulo tem como objetivo apresentar um mapeamento sistemático das abordagens tecnológicas para a análise semântica de mapas conceituais. O mapeamento foi elaborado a partir de uma revisão bibliográfica da área, entre os anos de 2015 e 2020. Posteriormente, aplicamos um filtro nessas abordagens para selecionar os trabalhos relacionados de interesse para nosso estudo e discutir suas características, permitindo uma análise dos resultados alcançados e elicitando de novos requisitos para trabalhos futuros.*

*Este capítulo está organizado da seguinte forma: a Seção 3.1 apresenta o processo da revisão da literatura; A Seção 3.2 apresenta um resumo dos principais resultados descobertos; A Seção 3.3 apresenta uma análise dos trabalhos relacionados; e a Seção 3.4 apresenta as considerações preliminares deste capítulo.*

## 3.1 O Processo da Revisão

Foi realizada uma revisão da literatura para mapear os estudos relacionados à extração de informações semânticas sobre a representação do conhecimento em mapas conceituais, a fim de buscar compreensão acerca do que o aluno representou ou o que ele sabe quando da construção do mapa, a partir da análise significativa das proposições construídas por ele.

Como o estado da arte da análise semântica de mapas conceituais não está de acordo com as diretrizes padrão, é difícil categorizar questões relacionadas. Este estudo visa fornecer um esquema de análise mais sistemático dos trabalhos neste contexto.

Esta revisão foi conduzida seguindo as diretrizes sugeridas por [Petersen et al. \(2008\)](#).

### 3.1.1 Questões de Pesquisa

A questão inicial que motivou essa categorização foi: Quais abordagens tecnológicas estão sendo desenvolvidas para a análise semântica de mapas conceituais? As seguintes questões de pesquisa foram definidas:

- (RQ1) – Quais são as principais características das abordagens tecnológicas neste contexto?

- (RQ2) – O que se sabe atualmente sobre os benefícios, desafios e limitações das abordagens?
- (RQ3) – Quais métodos e técnicas existem para apoiar o desenvolvimento dessas abordagens?
- (RQ4) – Quais as características e limitações das abordagens propostas para a avaliação semântica de mapas conceituais em Português e Inglês?
- (RQ5) – Qual ferramenta ou serviço da web está disponível para a avaliação semântica de mapas conceituais?

### 3.1.2 Pesquisa sobre os Estudos Primários

A partir dessas questões de pesquisa, definimos fontes de busca, bem como critérios de inclusão e exclusão. A estratégia de busca incluiu apenas bancos de dados eletrônicos, e são eles: IEEEExplore Digital Library, ACM Digital Library e Elsevier Science Direct.

A Scopus (Elsevier) foi incluída por ser uma das maiores fontes de referência, que incorpora trabalhos de diversas bases, tais como: ACM Digital Library, IEEE Xplore, ScienceDirect, entre outras. Porém, apesar de a Scopus englobar várias bases de publicações científicas, não existem garantias de que o resultado contenha todos os trabalhos existentes referentes à consulta processada. Sendo assim, a fim de apresentar maior completude e solidificação do conteúdo que pode ser adquirido, fez-se necessário a realização de consultas complementares nas bases IEEEExplore Digital Library e ACM Digital Library.

A estratégia utilizada para a construção da string de busca precisava envolver tanto palavras-chave que tivessem relação com mapas conceituais, quanto com processamento de linguagem natural e análise semântica, de forma que a string tivesse relação com as questões de pesquisa formuladas. Assim, nessas fontes de pesquisa, foram utilizadas as seguintes palavras-chave:

(“natural language processing” OR “nlp” OR “semantic analysis”) AND (“concept map” OR “concept mapping” OR “concept maps”)
--

Tabela 1 – *String* de busca.

Inicialmente, a seleção dos estudos potencialmente relevantes foi determinada pela análise do título, palavras-chave e resumo. Em seguida, a seleção dos estudos foi determinada pela leitura do artigo na íntegra.

Para a inclusão do estudo, foram considerados os seguintes critérios:

- (IC1) – Estudos escrito em Inglês ou Português.

- (IC2) – Estudos que abordam algumas das questões de pesquisa.
- (IC3) – As diferentes versões da obra publicadas por um autor na mesma abordagem.

Para a exclusão do estudo, foram considerados os seguintes critérios:

- (EC1) – Estudos repetidos. Se um estudo estiver disponível em mais de uma fonte de pesquisa, ele será considerado apenas na primeira vez que for encontrado.
- (EC2) – Estudos não científicos (notas, índice, editoriais, prefácios).
- (EC3) – Estudos irrelevantes para a pesquisa.
- (EC4) – Estudos cujos arquivos não puderam ser acessados pela instituição.

Depois de aplicar a string de pesquisa às fontes de pesquisa, 223 artigos foram retornados. Após o download, apenas 63 artigos foram considerados potencialmente relevantes na primeira seleção. Na segunda seleção, foi realizada uma melhor análise dos estudos primários, onde todos os artigos foram lidos e 19 artigos relevantes foram selecionados. A Tabela 1 resume o processo de seleção e apresenta a quantidade de trabalhos identificados em cada etapa.

Fonte	Estudos Recuperados	1ª Seleção	2ª Seleção				
			Irrelevante	Repetido	Não-Científico	Sem acesso	Estudo Primário
IEEE Xplore	90	25	7	7	0	0	11
ACM	83	24	11	3	5	0	5
Science Direct	50	14	5	1	0	5	3
<b>Total</b>	<b>223</b>	<b>63</b>	<b>23</b>	<b>11</b>	<b>5</b>	<b>5</b>	<b>19</b>

Tabela 2 – Processo de seleção do estudo primário.

A busca foi limitada a um determinado período, compreendido entre os anos de 2015 e 2020, ou seja, os últimos cinco anos. Tal período foi definido com a finalidade de encontrar trabalhos recém publicados ou que estejam em andamento nesse domínio. O gráfico da Figura 7 ilustra a concentração dos estudos por ano. Podemos observar que a maior concentração de estudos nesta área ocorreu nos anos de 2017 e 2019.



Figura 7 – Concentração de estudos por ano.

## 3.2 Análises dos Resultados

Esta seção traz um resumo dos principais resultados descobertos nesta pesquisa e analisa a sua importância para o domínio da Informática na Educação.

A fim de responder as questões da pesquisa, categorizamos inicialmente os estudos primários pelo modo que fizeram a avaliação da aprendizagem, se foi através da comparação entre mapas, de sistema de pontuação, de comparação e pontuação, ou do uso de semântica.

- (i) *Comparação entre mapas*: apuramos se a avaliação da aprendizagem foi feita por meio da comparação entre mapas. Na maior parte dos casos, se dá entre um mapa de referência e um mapa do aprendiz;
- (ii) *Sistema de pontuação*: averiguamos se a avaliação da aprendizagem foi feita a partir de um sistema de pontuação;
- (iii) *Comparação e pontuação*: investigamos se o estudo primário realizou a avaliação da aprendizagem utilizando tanto a comparação entre mapas quanto o sistema de pontuação;
- (iv) *Semântica*: perscrutamos se os estudos utilizaram algum tipo de análise semântica para avaliar os mapas conceituais.

Após essa primeira categorização, identificamos qual foi o meio utilizado pelos estudos primários para fazer a Análise Semântica dos Mapas Conceituais.

A seguir, na Figura 8, temos a tabulação dos dados encontrados nos 19 Estudos Primários selecionados. De acordo com as características definidas para a investigação das abordagens, é possível ter um conhecimento mais aprofundado dos estudos na área da avaliação da aprendizagem utilizando a semântica, através de mapas conceituais. A Seção 3.3 trata de algumas das soluções encontradas neste mapeamento sistemático.



Figura 8 – Tabulação dos dados referente aos trabalhos selecionados.

## 3.3 Resultados e Discussão

Este mapeamento sistemático teve por objetivo investigar soluções que envolvam análise semântica para realizar a avaliação da aprendizagem em mapas conceituais, como

forma de responder às questões de pesquisa definidas na Seção 2.1. Analisando os estudos encontrados neste mapeamento, foi possível encontrar algumas informações relevantes, que serão discorridas a seguir.

Muitos autores fizeram o estudo da semelhança semântica entre os conceitos de um grafo, e puderam lançar três grandes abordagens: abordagem baseada na distância, ou seja, na estrutura de um grafo, abordagem utilizando o conteúdo de informação dos conceitos e uma terceira que é o resultado da combinação das duas abordagens que a precedem (SALEM; CHENITI-BELCADHI; BRAHAM, 2017).

A abordagem de Salem, Cheniti-Belcadhi e Braham (2017) compara predicados e não conceitos, respeita a hierarquia durante o processo e recorre a um recurso eletrônico externo WordNet. A abordagem combina o uso dos mapas conceituais e as técnicas semânticas de similaridade. Já o trabalho de Vasques, Martins e Rezende (2018) utiliza como atributos a semântica contida nos verbos que constituem as relações existentes entre os conceitos para extrair relações implícitas no mapa conceitual. Esse tipo de análise semântica permite compreender as relações causais existentes entre os conceitos dos textos analisados. Assim, busca-se manter, na medida do possível, a carga semântica contextual, a fim de utilizar as relações extraídas na descoberta do conhecimento.

Em seu trabalho, Elsayed et al. (2019) considerou apenas um tipo de relação semântica, baseado em pré-requisito. Ou seja, um conceito C1 é pré-requisito de um conceito C2 quando C1 tem que ser cumprido e aprendido antes de C2.

O método desenvolvido por Fuentes-Lorenzo et al. (2019) visa adaptar métricas de similaridade baseadas em caminhos e de múltiplas fontes para trabalhar com a estrutura de categorização da Wikipedia como fonte de conhecimento, ao invés de estruturas lexicais tradicionais como o WordNet. Pois, segundo os autores, as informações obtidas de grandes corpora – incluindo a World Wide Web (WWW) – não são organizadas de forma específica e geralmente são necessárias técnicas de processamento de linguagem natural para obter resultados aceitáveis. Na pesquisa eles propuseram utilizar a Wikipédia, pois diziam que esta não possui tais limitações.

Manske e Hoppe (2017) introduziram tecnologias semânticas para extrair conceitos-chave de artefatos heterogêneos gerados pelo aluno. Artefatos de conhecimento são textos, mapas conceituais ou hipóteses. Já artefatos heterogêneos são definidos pelos autores como um conjunto de diferentes artefatos de conhecimento. Trata-se de uma abordagem que forma grupos heterogêneos. O mecanismo de formação de grupo semântico sugerido usa extração semântica do conteúdo gerado pelo aluno para criar grupos de aprendizagem heterogêneos em termos de diversidade de conhecimento. Em contraste com as antigas abordagens quantitativas baseadas em pontuação, isso segue a ideia de que um grupo se beneficia da complementaridade e diversidade de seus alunos. O trabalho desenvolvido por Boguski, Cury e Gava (2019) também faz a determinação de grupos semanticamente

semelhantes, nesse caso para sequenciar apresentação de conceitos mais adequados para cada aluno que serão passados através de um tutor, durante a construção do mapa conceitual. Para isso, é utilizada uma das técnicas existentes de mineração de dados.

O método de análise semântica proposto por [Hojas-Mazo et al. \(2019\)](#) baseia-se na extensão semântica de conceitos e num conjunto de regras para orientar os processos de integração, pesquisa e recuperação. Neste método, as informações semânticas associadas aos conceitos são capturadas do WordNet e um algoritmo de desambiguação de sentido de conceito é usado para reduzir a ambiguidade que pode surgir. A extensão semântica de conceitos é definida como o processo de associar a um conceito outros termos sinônimos identificados no WordNet.

A respeito do meio utilizado para fazer a Análise Semântica, a grande maioria das abordagens tecnológicas se utilizou de Medidas de Similaridade específicas, entre as diversas que existem, como pôde ser observado nos trabalhos de [Salem, Cheniti-Belcadhi e Braham \(2017\)](#), [Limongelli et al. \(2017\)](#) e [Elsayed et al. \(2019\)](#). Em seguida vieram abordagens que utilizaram o WordNet e Wikipedia, respectivamente. Como forma de sintetizar esta revisão, as Tabelas 3 e 4, evidenciam os diferentes grupos de acordo com suas características de meio de avaliação e técnica de análise semântica empregada, respectivamente.

Meio de avaliação da aprendizagem	Quantidade
Comparação entre mapas	7
Sistema de pontuação	2
Comparação + Sistema de Pontuação	1
Semântica	9

Tabela 3 – Meios de avaliação utilizados pelos estudos selecionados.

Técnica utilizada para a análise semântica	Quantidade
Medidas de Similaridade	11
Processamento de Linguagem Natural	1
Extração de Informações	1
WordNet	3
Wikipedia	2
Redes Neurais	1

Tabela 4 – Técnicas de análise semântica empregadas nos estudos selecionados.

### 3.4 Considerações sobre o capítulo

Neste capítulo, fizemos um mapeamento sistemático para conhecer e entender o funcionamento das abordagens tecnológicas que realizam a análise semântica em mapas

conceituais. Após a leitura contínua de diversos estudos encontrados nas três bases de dados (IEEE Xplore, ACM e Elsevier Science Direct) entre 2015 e 2020, observamos que a análise semântica é um assunto pouco explorado pelos pesquisadores, principalmente quando especificamos a área de interesse para o contexto de mapas conceituais. A questão inicial que motivou toda esta pesquisa progrediu-se em outras cinco questões e todas foram respondidas positivamente ao longo do trabalho.

Essa revisão de literatura agregou informações importantes para a identificação, compreensão, desenvolvimento e análise das abordagens de análise semântica em mapas conceituais. As técnicas utilizadas pelos estudos analisados continuam em uso nos dias de hoje, como as medidas de similaridade entre sentenças e o processamento de linguagem natural. Embora todas as abordagens usem análise semântica para avaliação de mapas conceituais, nenhuma delas faz uso das implicações significantes definidas por Piaget como forma de medição do aprendizado. Notamos também que nenhum dos trabalhos selecionados faz uso de mais de uma técnica de análise semântica, da mesma forma que nenhum dos estudos utiliza a semântica concomitantemente com outro meio de avaliação da aprendizagem.

Como forma de desenvolver uma técnica que supere algumas das falhas identificadas e inovar, buscamos usar um conjunto técnicas para promover uma análise semântica ainda mais aprofundada, com a utilização de processamento de linguagem natural, modelos treinados por redes neurais e medidas de similaridade. Do mesmo modo, desenvolvemos um sistema de avaliação da aprendizagem que avalie os mapas conceituais combinando semântica, sistema de pontuação e comparação de mapas.

A realização de um mapeamento objetivo para identificar e compreender abordagens neste contexto é um primeiro passo para melhorar e expandir a investigação na área. A carência de métricas torna inviável a comparação e avaliação da eficácia e cobertura das propostas. A partir de agora, nosso mapeamento já está apoiando nosso próximo passo, que consiste na elaboração de um modelo conceitual para uma nova abordagem tecnológica objetivando a análise semântica de mapas conceituais.

## CAPÍTULO 4

## O MODELO CONCEITUAL

*O objetivo desta pesquisa é “desenvolver uma arquitetura computacional para identificar as implicações significantes de Piaget aplicando análise semântica”, conforme apresentado no Capítulo 1. Neste capítulo apresentaremos a proposta de um modelo conceitual que recebe um mapa conceitual em inglês ou português, e retorna o mapa conceitual representado pelas implicações significantes e sua respectiva avaliação. As proposições do mapa conceitual são classificadas em: implicações locais, sistêmicas e estruturais.*

*Este capítulo está organizado da seguinte forma: a Seção 4.1 apresenta a visão geral do modelo conceitual; A Seção 4.2 descreve a caracterização adotada para a identificação das implicações significantes utilizando a análise semântica; A Seção 4.3 apresenta a etapa após a identificação das implicações; A Seção 4.4 apresenta o método para avaliação do mapa conceitual do aluno; A Seção 4.5 apresenta algumas considerações sobre o capítulo.*

## 4.1 O Modelo

O modelo proposto é estruturado a fim de analisar, de modo progressivo, as implicações representadas em cada mapa conceitual. A dimensão semântica é obtida com a utilização de técnicas de processamento de linguagem natural (PLN) de previsão e casamento de palavras em uma frase. Apesar de o modelo manipular mapas conceituais, todo o tratamento é feito por meio de relações matemáticas, com base no código JSON referente ao mapa conceitual. O modelo conceitual é apresentado na Figura 9 a seguir.

O modelo é composto por dois atores, aluno e professor, e dois componentes, repositório de mapas e repositório de mapas com implicações. O aluno envia o seu mapa conceitual para o Repositório de Mapas Conceituais, local em que são salvos os mapas da turma. Quando o repositório estiver completo, o Agente Implicante fará a identificação das implicações significantes de cada mapa conceitual. O professor também envia o mapa de referência para o Repositório de Mapas, para que este mapa seja utilizado como modelo para a correção dos mapas da turma. Como o mapa de referência marca o estabelecimento da construção do contexto semântico, é importante que o professor revise o mapa com bastante cautela antes de submetê-lo, para evitar que erros ou conceitos incorretos sejam inseridos de forma indevida.

O primeiro passo do modelo é procurar por implicações locais. Para isto, fazemos uma análise sintática de cada proposição do mapa para saber se ela contém verbo de

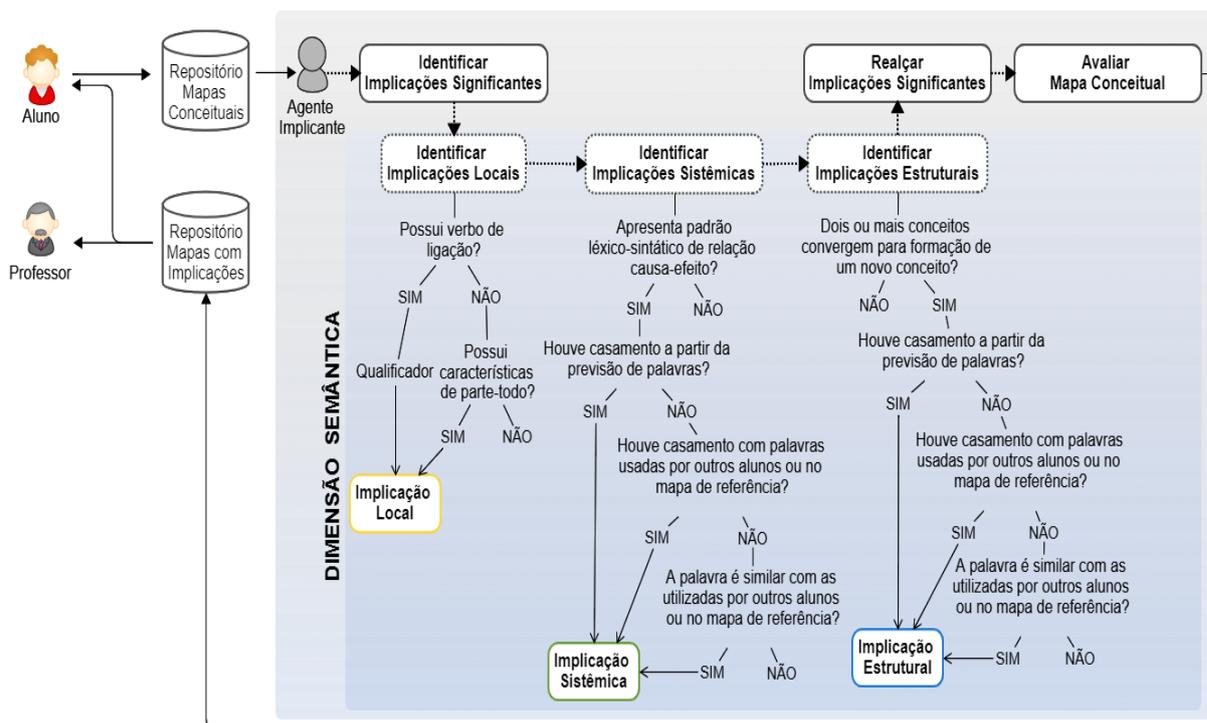


Figura 9 – Modelo conceitual.

ligação. Se houver, significa que há uma relação de qualificação, de forma que o segundo conceito da proposição caracteriza o primeiro, ocorrendo uma implicação local. Caso não haja verbo de ligação na relação buscamos encontrar características de parte-todo entre os dois conceitos da proposição. O objetivo é encontrar uma relação entre palavras em que o significado de uma (holônimo) refere-se a um todo do qual a outra (merônimo) é parte constituinte, como por exemplo “Carro x Para-brisa”, “Casa x Quarto” e “Rosto x Nariz”. Quando essas relações são encontradas, também efetuam-se implicações locais.

Em seguida, o Agente Implicante procura pelas implicações sistêmicas. Essas relações precisam apresentar um padrão léxico-sintático de causa-efeito. Para ocorrer, é necessário avaliar a proposição semanticamente para descobrir se ela está dentro do escopo considerado. Caso haja uma correspondência entre a palavra do conceito derivado da relação com alguma das palavras previstas, valida-se a implicação sistêmica. Caso não haja esse casamento, verificamos se há uma correspondência com palavras usadas por outros alunos ou no mapa de referência. Caso haja essa igualdade, há uma implicação sistêmica. Em caso negativo, o último teste é apurar se esse conceito utilizado pelo aluno possui similaridade com conceitos utilizados por outros alunos ou que estejam no mapa de referência. Se houver uma similaridade acima de 60%, consideramos que também há uma implicação sistêmica.

A última fase de identificar implicações significativas é identificar as estruturais. Ela inicia com a procura de dois ou mais conceitos que convergem para a formação de um novo conceito no mapa conceitual utilizando relações iguais. Esse requisito é essencial

para a formação de implicações estruturais. Novamente, é preciso avaliar as proposições semanticamente para descobrir se elas estão dentro do escopo considerado. Caso haja uma correspondência entre a palavra do conceito derivado da relação com alguma das palavras previstas, valida-se a implicação estrutural. Caso não haja esse casamento, verificamos se há uma correspondência com palavras usadas por outros alunos ou no mapa de referência. Caso haja essa igualdade, há uma implicação estrutural. Em caso negativo, o último teste é ver se esses conceitos utilizados pelo aluno possuem similaridade com conceitos utilizados por outros alunos ou que estejam no mapa de referência. Se houver uma similaridade acima de 60%, consideramos que também há uma implicação estrutural.

Após isso, as implicações significantes identificadas são realçadas por meio de cores. E por fim é feita a avaliação do mapa conceitual, de forma a atribuir uma nota a ele. A nota é formada pelo somatório da nota de todas as proposições do mapa. Todo esse processo é detalhado na Seção 4.4. Por fim, o mapa conceitual examinado é enviado para o repositório de mapas com implicações, que armazena todos aqueles que foram analisados.

Os estudantes poderão acessar o modelo a qualquer instante para saber como está a qualidade de seu mapa. Por exemplo, se somente amarelo, mostrando apenas implicações locais, isto indicará ao aluno a necessidade de prosseguir até um nível superior, chegando às implicações estruturais. Já o professor, ao observar os mapas coloridos com as implicações, pode perceber em que nível um aprendiz e sua turma alcançaram em determinado momento.

A verificação de semântica das implicações sistêmicas e estruturais, por meio de previsão de palavras, é realizada em dois níveis. No primeiro nível, realizamos uma verificação por meio de um modelo de redes neurais pré-treinado com bilhões de textos contidos em livros e na Wikipédia de forma a permitir o reconhecimento de amarrações semânticas. A previsão de palavras é realizada por um algoritmo chamado BERT (Bidirectional Encoder Representations from Transformers). É um modelo de representação de linguagem proposto por pesquisadores do Google AI Language que consiste em um codificador Transformer multicamada bidirecional (DEVLIN et al., 2018). Ao contrário do modelo de linguagem tradicional, que prevê uma palavra a partir de seu contexto esquerdo, ou seja, a partir de palavras anteriores a ela, esse modelo prevê uma palavra a partir de seu contexto esquerdo e direito. Isso significa que ele também prevê palavras considerando vocábulos escritos à direita da palavra a ser prevista. Resultados mostraram (DEVLIN et al., 2018) que um modelo de linguagem treinado bidirecionalmente pode ter um senso mais profundo de contexto e fluxo de linguagem do que modelos de linguagem de direção única. Para o corpus de pré-treinamento do BERT foram utilizados o BooksCorpus (800 milhões de palavras) e a Wikipédia em inglês (2.500 milhões de palavras). Esse tipo de modelo consegue prever com precisão relações contextuais entre palavras.

No segundo nível, comparamos se a palavra predita se encontra também no mapa

referência e nos mapas conceituais dos outros alunos. Essa abordagem, objetiva evidenciar um fechamento contextual, ampliando, dessa forma, a dimensão semântica, uma vez que, se o conceito a ser predito também aparecer nos outros mapas com um suporte e confiança singular, isso indicará que ele é compreensível por toda a turma de alunos, conjuntamente (BOGUSKI; CURY, 2018). Neste cenário, caso haja um casamento de palavras, por consequência, haverá uma implicação sistêmica ou estrutural.

Além disso, se o conceito estiver em todos os mapas conceituais da turma, mas não estiver no mapa de referência elaborado pelo professor, pode significar alguma divergência e o professor poderá verificar o porquê de a turma ter considerado tal conceito.

Contudo, em alguns casos, a palavra buscada não possui correspondência satisfatória em nenhuma das situações anteriores. Para isto, utilizamos um modelo baseado em Continuous Bag of Words (CBOW): o Word2Vec (“Word to Vector”). Ele permite encontrar a similaridade semântica entre duas palavras, mesmo que sejam diferentes. Esse modelo de processamento de linguagem natural transforma palavras em vetores dimensionais de características a fim de serem comparadas, dessa forma, diferentes sentidos da palavra são combinados em um único vetor. O modelo Word2Vec gera correspondências a partir da comparação de diferentes características, possibilitando uma abordagem de grande eficácia para verificação de similaridade semântica (MIKOLOV et al., 2013).

Essa estratégia procura avaliar a similaridade semântica entre a palavra que o aluno utilizou e demais palavras contidas nos mapas do repositório de mapas com implicações. Dessa forma, o Word2Vec foi aplicado no segundo nível de análise das implicações sistêmicas e estruturais, já que ele utiliza apenas duas palavras, que serão comparadas, para encontrar a similaridade entre elas. Já o modelo BERT, citado anteriormente, gera correspondências que nos permitem ter várias representações vetoriais para a mesma palavra, com base no contexto e sentido em que ela foi usada, necessitando de uma sentença ou frase para obter melhores resultados.

Neste nível é feita uma representação vetorial para cada palavra, e quanto maior for o valor entre duas palavras, mais similares elas são. Se as palavras tiverem um alto nível de similaridade, acima de 60%, significa que aquela proposição é válida e também representa uma implicação sistêmica ou estrutural. Todos os processos feitos tanto pelo spaCy quanto pelo BERT, serão detalhados no Capítulo 5.

## 4.2 Análise semântica em mapas conceituais

A análise semântica de um mapa conceitual representa um papel importante na tradução do conhecimento humano, na forma de mapas conceituais, em representações rigorosas e inequívocas para processamento posterior por computadores (HAO; YAN; CHI-WAI, 2010). Por meio da análise semântica, é possível extrair informações mais

precisas a partir de uma representação do conhecimento e a compreensão de um indivíduo acerca do que ele representou ou sobre o que ele sabia no momento da construção do mapa.

Nesta seção apresentamos a lógica por trás da identificação das implicações significativas, bem como a estratégia utilizada para análise semântica das proposições. Utilizaremos o mapa de referência com implicações da Figura 10 como base para os nossos exemplos.

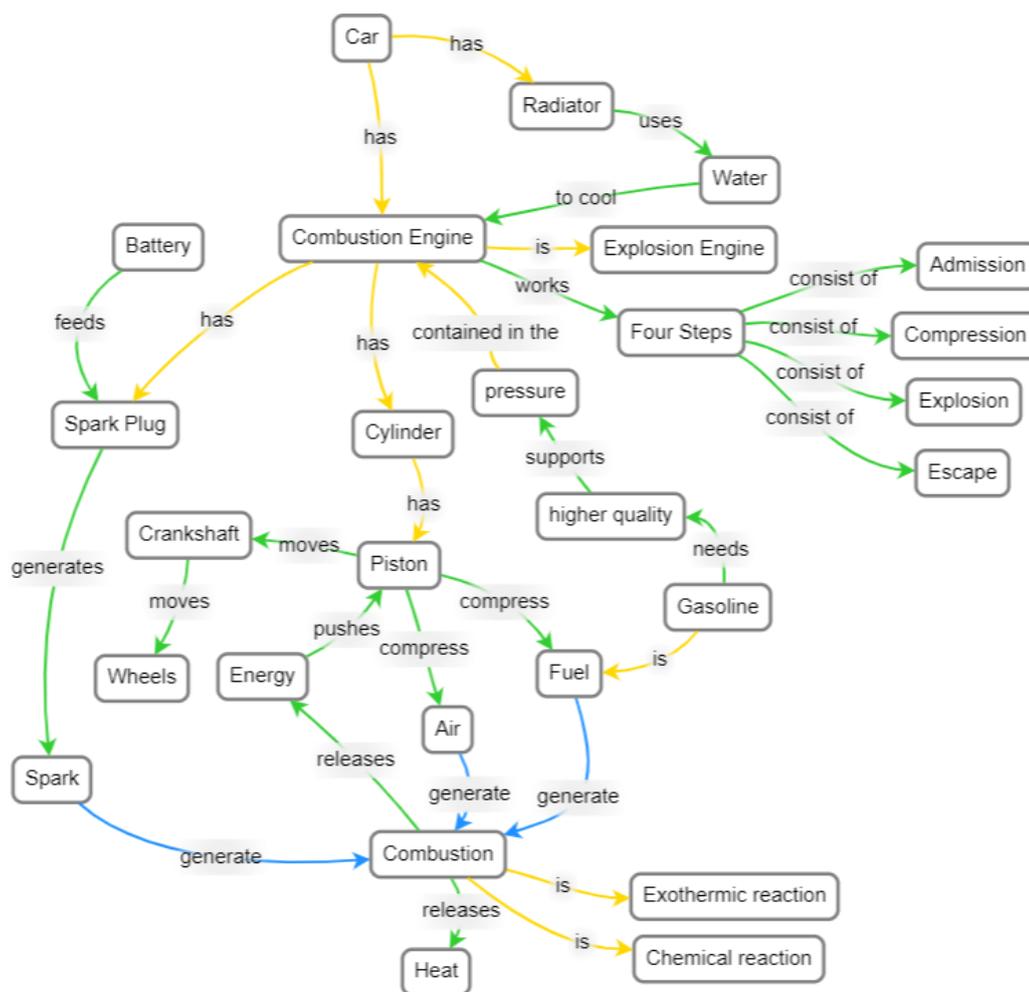


Figura 10 – Mapa de referência com avaliação das implicações significativas.

Fonte: Elaborado pelo autor.

### 4.2.1 Identificar Implicações locais

Para identificar implicações locais, procuramos verificar proposições cujo segundo conceito qualifica o primeiro. Neste contexto, quando um aluno utiliza um verbo de ligação, tal como ser/estar, parecer e tornar-se, a proposição passa a conter um predicativo do sujeito, o qual possui função de atribuir qualidades ao sujeito. Em relações que contenham tais verbos, é possível dizer que, semanticamente, o aluno entende que um conceito é um qualificador de outro, de forma a dar novas características a este. Um exemplo contido no mapa de referência da Figura 10 está na proposição “Combustão é uma reação química”.

Caso não haja um verbo de ligação na relação, utilizamos uma segunda abordagem.

Nestes casos, analisamos se o aluno está querendo representar características de parte-todo. Para isso, verificamos se há presença de verbos como *ter*, *pegar*, *poder*, *conter*, *possuir*. Esses verbos também caracterizam as implicações locais, porém indicam outro tipo de relação semântica, palavras que designam uma parte de um todo. Como exemplo na Figura 10, temos “Carro tem radiador”.

Em relação ao modelo anterior desenvolvido por [Moreira \(2019\)](#), cuja análise era feita por uma base de verbos definidos como "locais", as validações são feitas por meio de uma análise sintática da proposição, para a identificação de predicativos do sujeito, que semanticamente denotam a ideia de qualificação de um sujeito.

#### 4.2.2 Identificar Implicações sistêmicas

No domínio das implicações sistêmicas, primeiro é utilizado um conjunto de padrões léxico-sintáticos definidos por [Sorgente, Vettigli e Mele \(2013\)](#) que representam a estrutura de relações causais em sentenças, seguido por uma validação semântica. Os padrões sintáticos a serem identificados são:

- Os verbos causais simples são verbos únicos que têm o significado de “ação causal” (por exemplo, *gerar*, *disparar*, *criar* e assim por diante);
- Os verbos frasais são frases que consistem em um verbo seguido por uma partícula (por exemplo: *resulta em*);
- Substantivos + preposição são expressões compostas por um substantivo seguido por uma preposição (por exemplo: *causa de*);
- Os verbos causativos passivos são verbos em voz passiva seguidos da preposição *por* (por exemplo: *causado por*, *acionado por* e assim por diante);
- Preposições simples são preposições que podem ser usadas para ligar “causa” e “efeito” (por exemplo: *de*, *depois* e assim por diante).

Após essa primeira verificação léxico-sintática, aplica-se a mesma lógica da análise semântica para as implicações estruturais, a qual está descrita na Seção 4.2.4. Como exemplo de implicação sistêmica na Figura 10, temos a proposição “Combustão libera calor”.

No modelo anterior desenvolvido por [Moreira \(2019\)](#), a análise sintática das implicações sistêmicas consistia em identificar se os verbos das relações estavam contidos em uma base de verbos transitivos diretos criada manualmente. Desse modo, não era considerado o contexto da frase em que o verbo estava inserido para definir sua transitividade. O modelo atual passou a realizar a análise sintática com o apoio de uma biblioteca de processamento avançado de linguagem natural, abolindo o uso das base de verbos. A análise semântica,

que não existia anteriormente, foi efetuada com auxílio de um modelo treinado por redes neurais, por medidas de similaridade e por casamento de palavras.

### 4.2.3 Identificar Implicações estruturais

Em relação às implicações estruturais, primeiro é feita uma vistoria na topologia do mapa conceitual, buscando ocorrências de proposições em que vários momentos conceituais convergem para formação de um outro conceito, ou seja, a junção de diferentes conceitos que implicam na formação de um novo por meio de um mesmo verbo ou locução verbal.

Por exemplo, quando se observa, a partir da topologia do mapa referência da Figura 10, que os conceitos “Combustível”, “Oxigênio” e “Faísca” concorrem para a construção do conceito “Combustão”, significa que tais proposições são possíveis implicações estruturais, mas precisam passar por uma validação semântica. No modelo predecessor desenvolvido por [Moreira \(2019\)](#) apenas verificava-se a topologia do mapa para encontrar conceitos que convergiam para a formação de um novo, sem qualquer tipo de validação semântica. Nesta pesquisa, o novo modelo desenvolvido aplica a análise semântica com auxílio de um modelo treinado por redes neurais, medidas de similaridade e casamento de palavras. Essa novidade permite verificar aspectos relacionados ao significado das relações entre os conceitos.

### 4.2.4 Análise semântica de implicações

Para o nível de implicações sistêmicas e estruturais, precisamos saber se a proposição possui um valor semântico apropriado. Essa validação é feita a partir das técnicas de predição e casamento de palavras por meio da utilização de técnicas de PLN.

A técnica de predição consiste em prever possíveis palavras com base no contexto fornecido por uma sentença. Ao ocultar, por exemplo, o segundo conceito de uma proposição, é possível apontar eventuais palavras que se encaixam semanticamente no lugar do conceito “ocultado”. Com isso conseguimos ter uma ideia se a palavra utilizada pelo aluno no conceito resultante da proposição é adequada para o contexto fornecido pelo conceito e relação anteriores. Essa técnica não foi utilizada em nenhum dos estudos encontrados pelo mapeamento sistemático apresentados na Seção 3.3, e é realizada com o apoio do BERT, que será discutido na Seção 5.3. Trata-se de uma nova proposta que esta pesquisa apresenta para a comunidade científica e educação em geral.

A técnica de casamento de palavras e medidas de similaridade foi realizada com apoio de uma biblioteca de software de código aberto para processamento avançado de linguagem natural, e baseou-se nos trabalhos de [Salem, Cheniti-Belcadhi e Braham \(2017\)](#), [Limongelli et al. \(2017\)](#) e [Elsayed et al. \(2019\)](#).

O primeiro passo consiste em criar uma sentença contendo os conceitos de origem

com o verbo ou locução verbal da relação que os une, com o intuito de analisar sua semântica. Depois disso, é feita uma previsão de possíveis palavras que completam esta sentença naquele contexto semântico, buscando fazer um casamento de palavras.

Retomando o exemplo utilizado na Subseção 4.2.3, temos a seguinte sentença:

*“Combustível, Oxigênio e Faísca geram Combustão”*

A previsão de palavras é uma tarefa de processamento de linguagem natural que fornece uma frase parcial, onde as palavras mascaradas são identificadas para completar a frase a fim de produzir uma sentença que possua significado. Assim, o conceito derivado da proposição é mascarado (oculto) para que possam ser previstas e encontradas palavras que deem uma dimensão semântica à sentença.

No exemplo acima, o conceito “Combustão” é retirado da sentença e substituído pelo termo [MASK] para que se investigue se ele está contido no rol de palavras adequadas para o contexto dessa sentença. Caso ele seja encontrado, dizemos que houve um casamento de palavras. E, assim, estaremos diante de uma proposição com significação válida, e por conseguinte, uma implicação estrutural.

*“Combustível, Oxigênio e Faísca geram [MASK]”*

Ao analisar o contexto da frase, o algoritmo apresenta, por exemplo, 100 possíveis palavras cujos sentidos são possíveis de serem utilizados naquele campo semântico. Esse número de palavras é ajustado para que haja uma melhor precisão do modelo. Para o exemplo da proposição formulada acima, essas são algumas das palavras retornadas que se enquadram na posição do termo [MASK]:

[‘combustão’, ‘incêndios’, ‘reações’, ‘aquecimento’, ‘fogo’, ‘danos’, ‘acidentes’, ...]

Como podemos observar, a palavra “combustão” faz parte do rol de palavras que pertencem semanticamente ao contexto da proposição criada pelo professor no mapa referência da Figura 10. Logo, essa proposição é identificada como implicação estrutural.

### 4.3 Realçar Implicações Significantes

Após processar o mapa conceitual, a penúltima etapa dentro do agente implicante é realçar as implicações significantes, atribuindo a cor amarela para implicações locais, verde para sistêmicas e azul para estruturais.

Pode haver casos de relações que foram classificadas como locais que também se enquadram nos requisitos das implicações estruturais. Dessa forma, optou-se pela seguinte

ordem de precedência: Implicações Locais -> Implicações Estruturais -> Implicações Sistêmicas. Portanto, uma implicação caracterizada como Sistêmica, pode vir a se tornar uma implicação Estrutural posteriormente caso atenda aos requisitos, mas uma Implicação Local não.

## 4.4 Avaliação do Mapa Conceitual

Por fim, chegamos à etapa de avaliação do mapa conceitual. Sabemos que a avaliação é parte integrante do docente na atividade escolar e precisa ser percebida como um elemento importante no processo de formação do aluno, possibilitando uma reflexão constante das habilidades e dificuldades, bem como do papel do docente para as transformações na sociedade.

A avaliação da aprendizagem traz benefícios não só para os alunos, como também para os docentes. No caso dos estudantes, existe a possibilidade de verificar o andamento do seu aprendizado e buscar metodologias para impulsionar o seu desenvolvimento. Além disso, professores podem incentivar a autoavaliação nos alunos, e estimular a sua participação ativa na aprendizagem.

Para os professores, essa estratégia é uma oportunidade para verificar se os alunos e a turma conseguiram atingir as metas definidas. Dessa forma, é possível trazer um novo direcionamento às ações pedagógicas para que os objetivos sejam atingidos.

Essa etapa é uma das mais importantes do nosso modelo, pois é onde podemos de fato quantificar o conhecimento que o aluno soube expressar no mapa conceitual no momento em que ele o construiu, comparando com aquilo que o professor esperava que fosse representado.

O cálculo da nota do aluno tem como base o número de implicações significantes presentes no mapa de referência elaborado pelo professor. Além disso, é necessário que o aluno tenha formado implicações que contêm valor semântico igual ou próximo com o que era esperado pelo professor. Portanto, é preciso que o aluno tenha construído o número total de cada tipo de implicação significativa e alcance o nível de entendimento pretendido para obter a nota máxima em seu mapa conceitual.

Por exemplo, se um tema a ser abordado é sobre o funcionamento do motor a combustão, e o aluno crie relações que fujam da temática, então tais implicações recebem uma nota inferior por estarem fora do escopo, porém não deixam de ser identificadas no mapa conceitual. Ademais, é justo pontuar eventuais implicações que o aluno possa formular e que não estejam presentes no mapa de referência, seja porque o aluno representou algo que foi esquecido pelo professor ou porque foi além do que se esperava dele.

A avaliação do mapa conceitual consiste em verificar sintática e semanticamente

conceitos e seus relacionamentos. É necessário averiguar se eles foram construídos corretamente e se fazem parte do escopo semântico da atividade. A avaliação é feita para cada relação presente no mapa conceitual para todos os diferentes níveis de implicações significantes definidas por Piaget. Para isto, foram estabelecidos 7 critérios para auxiliar o processo de avaliação:

- **Critério 1 – Implicação válida:** A implicação foi caracterizada corretamente pelo aluno, possuindo um valor semântico próprio e sem qualquer insuficiência sintática.
- **Critério 2 – Implicação contida no mapa de referência:** Além das características do critério anterior, a implicação definida pelo aluno está dentro do escopo semântico da atividade e presente no mapa conceitual de referência.
- **Critério 3 – Implicação sem conceito de origem:** O conceito de destino e o relacionamento estão corretamente representados, mas o conceito de origem não. Isso pode indicar que o aluno sabe que existe uma relação entre dois conceitos, mas ainda desconhece um deles.
- **Critério 4 – Implicação sem conceito de destino:** O conceito de origem e o relacionamento estão corretamente representados, mas o conceito de destino não. Isso pode indicar que o aluno sabe que existe uma relação entre dois conceitos, mas ainda desconhece um deles.
- **Critério 5 – Implicação com relação não descrita:** Os conceitos e o relacionamento estão corretamente ligados, mas a não foi descrita. Isso pode indicar que o aluno sabe que existe uma relação entre os conceitos, mas ainda a desconhece.
- **Critério 6 – Implicação não possui verbo ou locução verbal:** A presença de um verbo ou locução verbal é condição necessária para que ocorra uma implicação segundo às classificações dadas por Piaget.
- **Critério 7 – Implicação inválida:** Esse critério ocorre quando a implicação não conseguiu ser validada semanticamente e/ou não recebeu nenhuma das classificações propostas por Piaget.

Na Figura 11 exibimos os critérios de pontuação para avaliação das relações e seu valor pode ser calculado simplesmente somando-se cumulativamente o valor de cada critério para o qual o conceito em questão é pertinente.

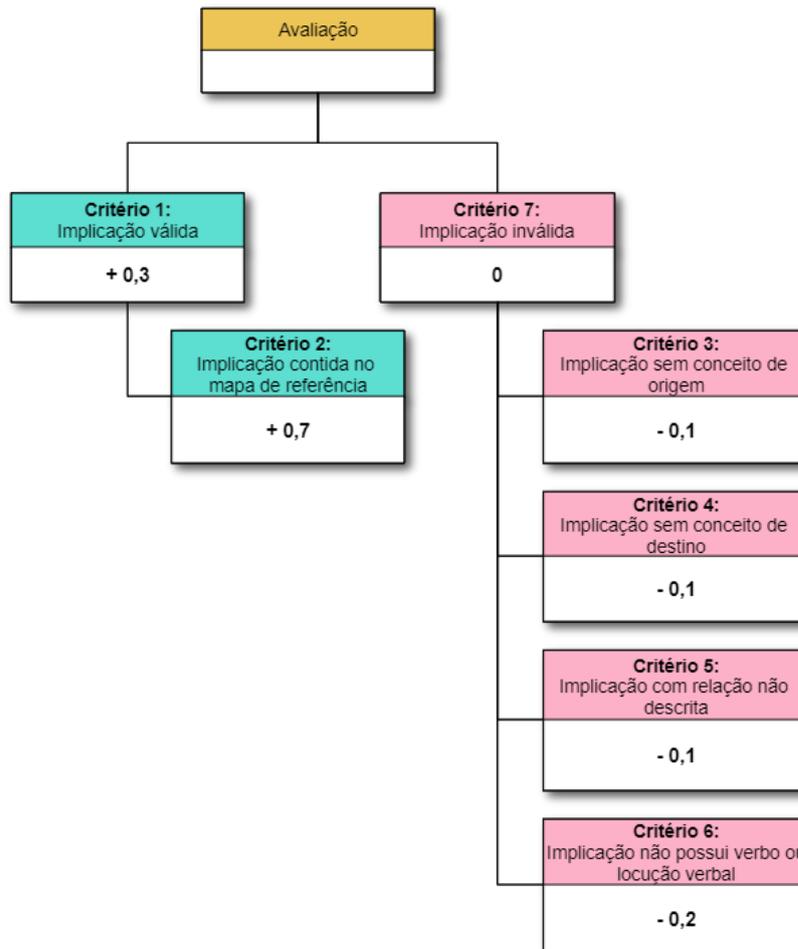


Figura 11 – Critérios de pontuação na avaliação do mapa conceitual.

Para exemplificação do processo avaliativo do mapa conceitual, considere os conceitos representados por um aluno, conforme a Figura 12 abaixo e o mapa referência na Figura 10 em que esses conceitos foram indicados.

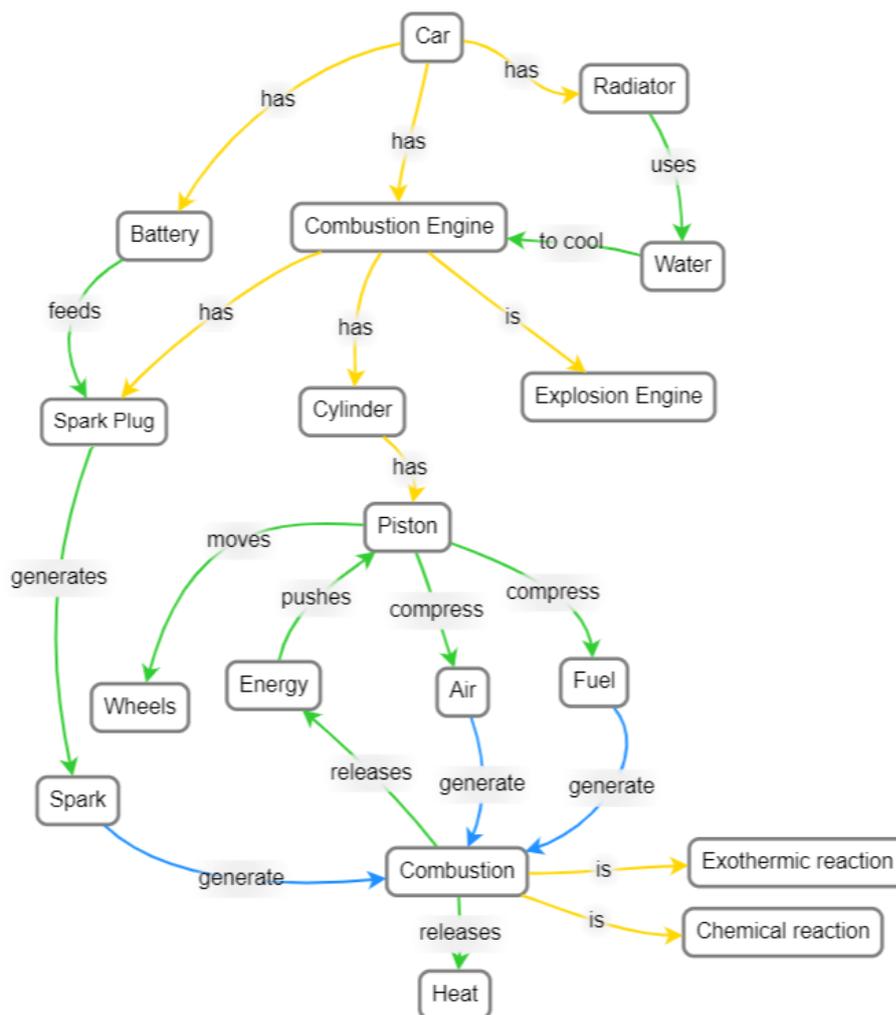


Figura 12 – Mapa conceitual de um aluno avaliado pelas implicações significantes.

Fonte: Elaborado pelo autor.

Classificando as implicações de acordo com os critérios mencionados anteriormente, identificamos que esse aluno representou 69% (6,9 pontos do total de 10 pontos possíveis) das implicações que eram esperadas pelo mapa de referência. Essa informação é obtida no código JSON retornado pelo modelo. No capítulo a seguir será explanada toda a informação que o modelo é capaz de fornecer.

Quando comparamos esse resultado com a quantidade de proposições pretendidas pelo professor em seu mapa de referência, podemos ter uma visão holística da avaliação. Na situação proposta, o professor definiu 32 proposições como relevantes para a tarefa, sendo que o aluno representou 22 desse total, obtendo a nota aproximada de 69%.

## 4.5 Considerações sobre o capítulo

Neste capítulo apresentamos o modelo conceitual da solução proposta para a identificação automática das implicações significantes definidas por Piaget em um mapa

conceitual utilizando análise semântica. De modo geral, este modelo apresenta uma evolução relevante em relação aos trabalhos anteriores e é um primeiro passo para a introdução da análise semântica de mapas conceituais.

As técnicas de PLN foram empregadas através do spaCy ([HONNIBAL; MONTANI, 2017](#)), uma biblioteca de código aberto gratuita para processamento de linguagem natural avançado em Python. Ela pode ser usada para construir sistemas de extração de informações ou de compreensão de linguagem natural, ou para pré-processar texto para aprendizado profundo. Na Seção 5.2 apresentamos mais informações sobre essa biblioteca. Já o BERT, explicado na Seção 5.3, é um algoritmo baseado em inteligência artificial capaz de entender muito melhor o significado e as relações entre palavras e frases e de reconhecer o contexto completo de cada palavra, identificando e levando em conta tanto os termos que vêm antes quanto os que vêm depois de cada vocábulo.

## CAPÍTULO 5

## MAPIMP 2.0

*Neste capítulo apresentamos o MAPimp 2.0, uma nova API Python que fornece um conjunto de bibliotecas de processamento de linguagem natural. Esta API é um dos resultados da pesquisa e foi motivada pela necessidade encontrada na implantação do modelo conceitual apresentado no Capítulo 4.*

*O capítulo está organizado da seguinte forma: a Seção 5.1 apresenta o MAPimp 2.0; As Seções 5.2 e 5.3 descrevem as bibliotecas spaCy e BERT; e a Seção 5.4 apresenta algumas considerações sobre o capítulo.*

## 5.1 Sobre o MAPimp 2.0

O MAPimp 2.0 é uma nova geração de um modelo computacional que auxilia na identificação de implicações significantes em mapas conceituais, disponível tanto para os professores quanto para os estudantes. A nova versão contém validações sintático-semânticas, que não existiam na versão anterior.

A partir do modelo, os professores poderão examinar de forma mais evidente e direcionada as implicações construídas pelos estudantes, permitindo um diagnóstico rápido sobre as deficiências do aprendizado. Já os estudantes poderão averiguar seus mapas e observar as implicações identificadas pela ferramenta, de forma a incrementar novas percepções sobre o seu próprio mapa. Ou seja, se o mapa de um estudante for colorido apenas com a cor amarela, então o professor e o aluno poderão analisar e constatar que o aprendizado está se desenvolvendo apenas em nível local, carecendo de prosseguir até um nível superior.

Trata-se de uma API em Python que identifica automaticamente as diversas formas de manifestação das implicações significantes em um mapa conceitual para os idiomas inglês e português. Esta API usa serviços comuns de Processamento de Linguagem Natural (PLN) de outros kits de ferramentas, como o spaCy, para fornecer tarefas mais específicas para inspeção de textos. O objetivo da solução é facilitar a análise de mapas conceituais para alunos e professores, proporcionando-lhes um feedback mais rápido. A construção do serviço se deu com base no trabalho desenvolvido por [Moreira, Aguiar e Cury \(2019\)](#).

Além disso, o modelo será integrado em um ambiente de aprendizagem assistida por tutores inteligentes sobre mapas conceituais ([BOGUSKI; CURY; GAVA, 2019](#)), a fim de que a identificação automática das implicações em um mapa conceitual produzido por um aluno possa permitir que um tutor interaja com ele, em tempo real, com base no conhecimento que ele foi capaz de expor naquele instante, para obter maiores indícios de aprendizado,

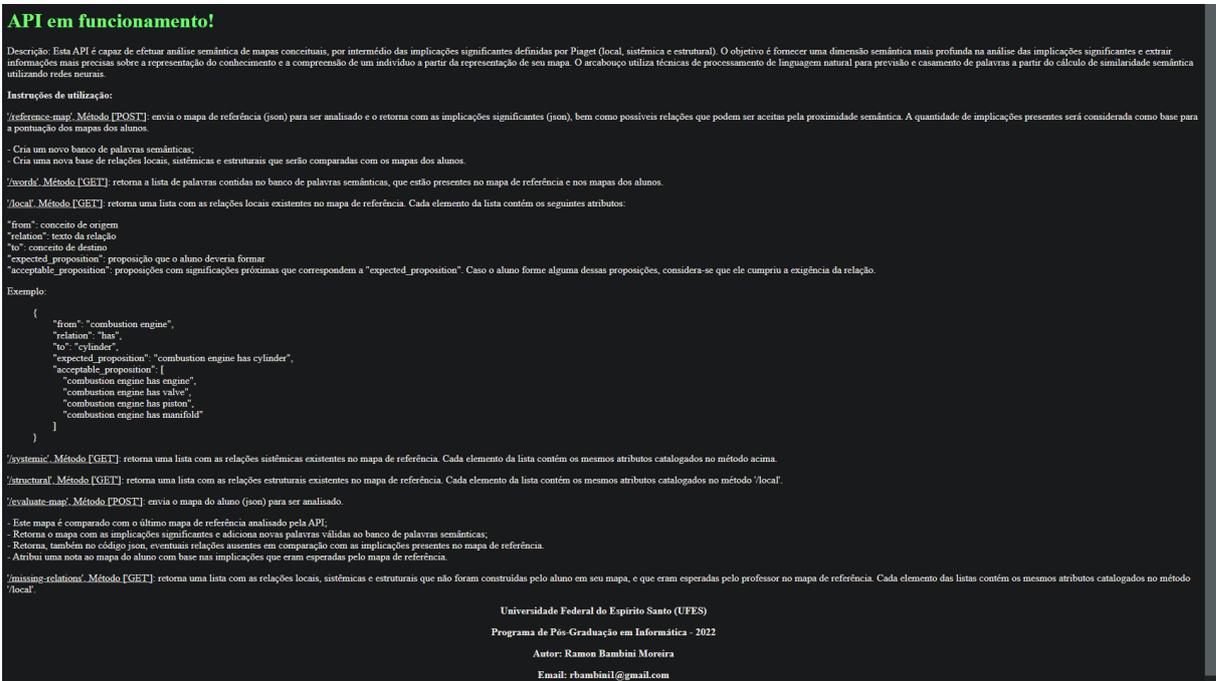
além da confirmação das implicações obtidas. Como consequência, o tutor pode sugerir novas ideias, novos conceitos, novos relacionamentos e levantar questionamentos ao aluno a partir daquilo que ele está representando no momento da construção do seu mapa. Essa cooperação trata-se por ser um trabalho correlato desenvolvido na mesma linha de pesquisa em nosso laboratório na Universidade Federal do Espírito Santo (UFES). A API está disponível como um serviço e em breve poderá ser usada por meio do padrão URI.

O serviço recebe como entrada um código JSON, referente ao mapa conceitual que será avaliado, e o retorna posteriormente com as implicações significantes identificadas através de cores, nas setas de ligação.

A API é composta por duas bibliotecas, SpaCy e Bert, que serão apresentadas neste capítulo nas seções 5.2 e 5.3, respectivamente. A seguir, demonstraremos todas as funcionalidades da API resultantes desta pesquisa.

### 5.1.1 Documentação da API

Uma página web principal foi criada contendo as informações de todos os métodos integrados à API. Dessa forma, quando um usuário acessa a URI da API em seu navegador, ele verá as informações constantes na Figura 13.



**API em funcionamento!**

Descrição: Esta API é capaz de efetuar análise semântica de mapas conceituais, por intermédio das implicações significantes definidas por Piaget (local, sistêmica e estrutural). O objetivo é fornecer uma dimensão semântica mais profunda na análise das implicações significantes e extrair informações mais precisas sobre a representação do conhecimento e a compreensão de um indivíduo a partir da representação de seu mapa. O arcabouço utiliza técnicas de processamento de linguagem natural para previsão e casamento de palavras a partir do cálculo de similaridade semântica utilizando redes neurais.

Instruções de utilização:

- `/reference-map`. Método [POST] envia o mapa de referência (json) para ser analisado e o retorna com as implicações significantes (json), bem como possíveis relações que podem ser aceitas pela proximidade semântica. A quantidade de implicações presentes será considerada como base para a pontuação dos mapas dos alunos.
- Cria um novo banco de palavras semânticas;
- Cria uma nova base de relações locais, sistêmicas e estruturais que serão comparadas com os mapas dos alunos.
- `/words`. Método [GET] retorna a lista de palavras contidas no banco de palavras semânticas, que estão presentes no mapa de referência e nos mapas dos alunos.
- `/local`. Método [GET] retorna uma lista com as relações locais existentes no mapa de referência. Cada elemento da lista contém os seguintes atributos:
  - `"from"`: conceito de origem
  - `"relation"`: tipo da relação
  - `"to"`: conceito de destino
  - `"expected_proposition"`: proposição que o aluno deveria formar
  - `"acceptable_proposition"`: proposições com significações próximas que correspondem a "expected\_proposition". Caso o aluno forme alguma dessas proposições, considera-se que ele cumpriu a exigência da relação.

Exemplo:

```
{
  "from": "combustion engine",
  "relation": "has",
  "to": "cylinder",
  "expected_proposition": "combustion engine has cylinder",
  "acceptable_proposition": [
    "combustion engine has engine",
    "combustion engine has valve",
    "combustion engine has piston",
    "combustion engine has manifold"
  ]
}
```

- `/systemic`. Método [GET] retorna uma lista com as relações sistêmicas existentes no mapa de referência. Cada elemento da lista contém os mesmos atributos catalogados no método acima.
- `/structural`. Método [GET] retorna uma lista com as relações estruturais existentes no mapa de referência. Cada elemento da lista contém os mesmos atributos catalogados no método `/local`.
- `/evaluate-map`. Método [POST] envia o mapa do aluno (json) para ser analisado.
  - Este mapa é comparado com o último mapa de referência analisado pela API;
  - Retorna o mapa com as implicações significantes e adiciona novas palavras válidas ao banco de palavras semânticas;
  - Retorna, também no código json, eventuais relações ausentes em comparação com as implicações presentes no mapa de referência.
  - Atribui uma nota ao mapa do aluno com base nas implicações que eram esperadas pelo mapa de referência.
- `/missing-relations`. Método [GET] retorna uma lista com as relações locais, sistêmicas e estruturais que não foram construídas pelo aluno em seu mapa, e que eram esperadas pelo professor no mapa de referência. Cada elemento das listas contém os mesmos atributos catalogados no método `/local`.

Universidade Federal do Espírito Santo (UFES)  
 Programa de Pós-Graduação em Informática - 2022  
 Autor: Ramon Bambini Moreira  
 Email: rbambini@gmail.com

Figura 13 – Tela principal da API.

Fonte: Elaborado pelo autor.

A API utiliza dois métodos do protocolo HTTP (Hypertext Transfer Protocol), que permite a comunicação entre clientes e servidores. Ou seja, o HTTP funciona como um protocolo de solicitação-resposta entre um cliente e um servidor.

- **GET**: é usado para solicitar dados de um recurso especificado (buscar informações).
- **POST**: é usado para enviar dados a um servidor para criar/atualizar um recurso (input de informações).

O primeiro recurso chamado *‘/reference-map’* é utilizado para enviar e salvar o mapa de referência, no formato JSON. Ele também passa por uma análise e o recurso retorna o mapa com as implicações significantes. A partir dele é criado um novo banco de palavras semânticas, o qual servirá como base para comparação com os conceitos dos alunos, bem como possíveis relações que podem ser aceitas pela proximidade semântica. Esse método também cria uma nova base de dados com as relações locais, sistêmicas e estruturais que serão usadas como parâmetro para comparação com os mapas dos alunos.

Como forma de fornecer ainda mais informações, a API ainda apresenta as seguintes informações, no código JSON, a respeito do mapa de referência:

- Lista de relações não contidas em nenhuma das três classificações de Piaget.
- Listas das implicações locais, sistêmicas e estruturais com informações relevantes para cada relação presente mapa. Essas informações incluem:
  1. Proposição esperada;
  2. Proposições aceitáveis, com base na proximidade semântica da proposição esperada.
  3. Conceito de origem;
  4. Conceito de destino;
  5. Conteúdo da relação.
- Lista com estatísticas do mapa de referência:
  1. Número de relações não identificadas entre as três classificações de Piaget.
  2. Número de implicações totais;
  3. Número de implicações locais.
  4. Número de implicações sistêmicas;
  5. Número de implicações estruturais.

Listagem 5.1 – Exemplo de informações extraídas do mapa conceitual de referência.

```
1 {
2   "implications": {
3     "local": [
4       {
5         "acceptable_proposition": [
```

```

6           "combustion engine has engine ",
7           "combustion engine has valve ",
8           "combustion engine has piston ",
9           "combustion engine has manifold "
10        ],
11        "expected_proposition": "combustion engine has cylinder ",
12        "from": "combustion engine ",
13        "relation": "has ",
14        "to": "cylinder "
15    }
16    "noImplication": []
17  }
18
19  "stats": {
20    "noImplications": 0,
21    "numberImplications": 32,
22    "numberLocal": 11,
23    "numberStructural": 3,
24    "numberSystemic": 18
25  }
26
27  ...

```

O recurso `/words` retorna a lista de palavras contidas no banco de palavras semânticas, que estão presentes no mapa de referência e nos mapas dos alunos.

O recurso `/local` retorna uma lista com as relações locais existentes no mapa de referência. Cada elemento da lista contém os seguintes atributos:

- `“from”`: conceito de origem;
- `“relation”`: texto da relação;
- `“to”`: conceito de destino;
- `“expected_proposition”`: proposição que o aluno deveria formar;
- `“acceptable_proposition”`: proposições com significações próximas que correspondem a `“expected_proposition”`. Caso o aluno forme alguma dessas proposições, considera-se que ele cumpriu a exigência da relação.

Os recursos `/systemic` e `/structural` têm a mesma finalidade, e retornam as suas respectivas listas.

O recurso `/evaluate-map` é utilizado para analisar o mapa conceitual do(s) aluno(s). O mapa é enviado, no formato JSON, para ser comparado com o último mapa de referência recebido pela API. Ele retorna o mapa com as implicações significantes; adiciona novas palavras no banco de palavras semânticas, caso ainda não estejam; também retorna, no código json do mapa, eventuais relações ausentes se comparadas com as implicações presentes no mapa de referência; e, atribui uma nota, em porcentagem, ao mapa do aluno

com base nas implicações que eram esperadas pelo mapa de referência, porém sem deixar de considerar outras relações que ele construiu.

O código em JSON enviado ao serviço precisa conter duas classes:

1. **nodeDataArray**: essa classe é um array que contém todos os conceitos do mapa conceitual. Cada conceito é listado separadamente como um objeto dentro do `nodeDataArray`. De modo a padronizar as informações, cada objeto do array deve ter, necessariamente, os seguintes atributos e seus respectivos tipos:

- a) **text**: contém as palavras, do tipo `String`, que representam os conceitos e relações;
- b) **loc**: contém as coordenadas para a localização exata do conceito na visualização gráfica do mapa;

Outros atributos podem aparecer dentro da classe `nodeDataArray`, porém não serão levados em conta para a análise das implicações.

2. **linkDataArray**: essa classe é um array que contém os atributos responsáveis por identificar as setas direcionadas nos mapas, ou seja, as setas que ligam conceito-relação e relação-conceito. Cada seta direcionada é representada através de um objeto dentro do array. Para cada objeto são necessários os seguintes atributos:

- a) **from**: identificador que utiliza o atributo `key` da classe `nodeDataArray`, para caracterizar onde a seta direcionada tem início;
- b) **to**: identificador que utiliza o atributo `key` da classe `nodeDataArray`, para caracterizar onde a seta direcionada tem fim.
- c) **text**: contém as palavras que cria a relação entre dois conceitos, formando uma proposição.

Outros atributos podem aparecer dentro da classe `linkDataArray`, porém não serão levados em conta para a análise das implicações. Posteriormente será adicionado o atributo `color` dentro de cada objeto na classe `linkDataArray` para identificar todas as implicações significantes.

Será utilizada a cor amarela para implicações locais, verde para sistêmicas e azul para estruturais. O que não for classificado como nenhuma das três implicações estruturais, retornará com a cor preto.

Assim, a classe `linkDataArray` posteriormente às implicações significantes ficaria parecido com o código exemplo abaixo:

Listagem 5.2 – Exemplo da identificação de uma relação estrutural no código json.

```
1 { ...
```

```

2 { "linkDataArray": [
3   { "from": "Fuel", "to": "Combustion", "text": "generate", "color": "Blue" }
4   ]
5 }

```

Neste caso, a relação foi identificada como sendo uma implicação estrutural.

Também, com o objetivo de extrair o máximo de informações, a API ainda apresenta as seguintes informações, no código JSON, à respeito do mapa do aluno:

- Lista de relações não contidas em nenhuma das três classificações de Piaget.
- Listas das implicações locais, sistêmicas e estruturais representadas no mapa do aluno. Para cada proposição é fornecida as informações abaixo:
  1. Lista contendo informações sobre a avaliação da proposição, entre elas:
    - a) Proposição esperada pelo professor, com base na proximidade semântica;
    - b) Proposição estabelecida pelo aluno;
    - c) Nota da proposição, com base nos critérios definidos na Seção 4.4;
    - d) Justificativa, também com base nos critérios definidos na Seção 4.4.
  2. Conceito de origem;
  3. Conceito de destino;
  4. Conteúdo da relação.
- Lista com estatísticas gerais do mapa do aluno:
  1. Número de Aproximações: proposições que foram identificadas a partir de aproximações semânticas
  2. Nota: dada em porcentagem a partir do cálculo da divisão do número total de implicações representadas pelo aluno pelo número de implicações totais do mapa de referência.
  3. Número de relações não identificadas entre as três classificações de Piaget.
  4. Número de implicações totais;
  5. Número de implicações locais.
  6. Número de implicações sistêmicas;
  7. Número de implicações estruturais.

Listagem 5.3 – Exemplo de informações extraídas do mapa conceitual do aluno.

```

1 {
2   "implications": {
3     "local": [

```

```
4      {
5          "evaluation": {
6              "expected_proposition": "combustion engine has cylinder",
7              "grade": 1.0,
8              "justification": "1;2",
9              "obtained_proposition": "combustion engine has cylinder "
10         },
11         "from": "combustion engine",
12         "relation": "has",
13         "to": "cylinder "
14     }
15     "noImplication": []
16 }
17
18 "stats": {
19     "aproximations": "3.33%",
20     "grade": "96.0%",
21     "noImplications": 1,
22     "numberImplications": 30,
23     "numberLocal": 9,
24     "numberStructural": 3,
25     "numberSystemic": 18
26 }
27
28 ...
```

O recurso `/missing-relations` retorna uma lista com as relações locais, sistêmicas e estruturais que não foram construídas pelo aluno em seu mapa, e que eram esperadas pelo professor no mapa de referência. Cada elemento das listas contém os mesmos atributos catalogados no método `/local`.

Já o recurso `/text-map` é um método de leitura do mapa conceitual que é capaz de gerar um texto a partir das proposições e topologia do mapa.

Por fim, o recurso `/implication` retorna o tipo de implicação significativa que a proposição enviada representa. É um recurso que fornece uma resposta mais veloz, útil principalmente para as ações que requeiram uma resposta em tempo real.

### 5.1.2 Exemplo de utilização

Para complementar a explicação do serviço, vamos demonstrar de forma prática, como fazer requisições à API. Para fazer uma requisição ao serviço, neste exemplo utilizaremos o Postman, uma aplicação que permite realizar requisições HTTP a partir de uma interface simples e intuitiva.

No Postman, adicione o endereço em que o serviço está sendo executado no campo “Enter request URL”. Neste exemplo, vamos fazer uma requisição com o recurso `/avaliar-mapa` para analisar um mapa conceitual. Como o serviço retorna o código JSON modificado a partir de um código JSON de entrada, precisamos fazer uma requisição do tipo POST.

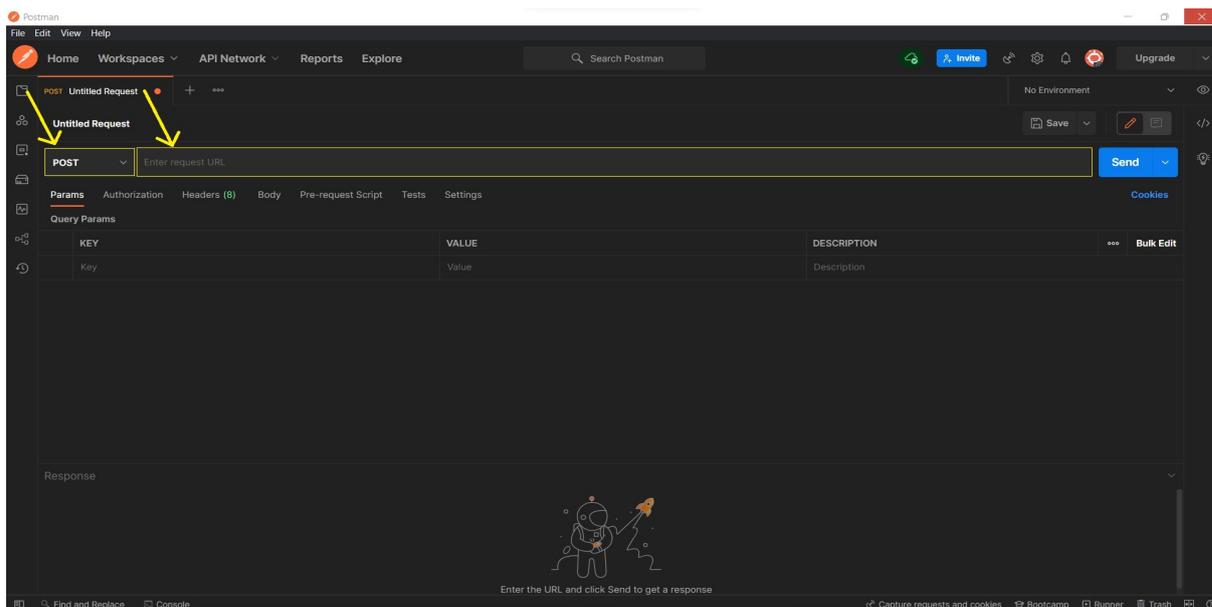


Figura 14 – Requisição do serviço com o método POST e o endereço da URL do serviço.  
Fonte: Elaborado pelo autor.

O próximo passo é inserir o código JSON a ser analisado pelo serviço. Para a demonstração desse exemplo, vamos utilizar o código json abaixo do mapa referente à Figura 5. Abaixo listamos alguns trechos referente a esse mapa conceitual:

Listagem 5.4 – Trecho do código json referente ao mapa conceitual da Figura 5.

```

1 { "class": "GraphLinksModel",
2   "nodeKeyProperty": "text",
3   "nodeDataArray": [
4     {"text": "Fuel", "loc": "-1299.1421062222919 -542.9233977623306"},
5     {"text": "Air", "loc": "-1367.3765369140629 -537.5688149189457"},
6     {"text": "Combustion", "loc": "-1395.1827291406262 -437.66497708691435"},
7     {"text": "Combustion Engine", "loc": "-1447.5804988662135 -796.4104308390025"},
8     {"text": "Chemical reaction", "loc": "-1231.247165532881 -403.8893546106598"},
9     {"text": "Cylinder", "loc": "-1443.021156061151 -693.7144465219516"},
10    {"text": "Energy", "loc": "-1483.4880952380947 -541.2676870748301"},
11    {"text": "Spark Plug", "loc": "-1578.9532311639712 -687.9783190420918"},
12    {"text": "Car", "loc": "-1386.4999334468064 -902.0656280517571"},
13    {"text": "Spark", "loc": "-1588.2551345014806 -477.1493848531477"},
14    {"text": "Gasoline", "loc": "-1172.5000122070305 -573.6625000000001"},
15    {"text": "Wheels", "loc": "-1558.231032080255 -529.005710524238"},
16    {"text": "Water", "loc": "-1177.5000488281248 -838.6125000000005"},
17    {"text": "Explosion Engine", "loc": "-1246.6937339529645 -790.7667855818598"}
18  ],
19 ],
20 "linkDataArray": [
21 {"from": "Combustion Engine", "to": "Cylinder", "text": "has"},
22 {"from": "Combustion Engine", "to": "Spark Plug", "text": "has"},
23 {"from": "Piston", "to": "Fuel", "text": "compress"},
24 {"from": "Piston", "to": "Air", "text": "compress"},
25 {"from": "Spark Plug", "to": "Spark", "text": "generates"},
26 {"from": "Air", "to": "Combustion", "text": "generate"},
27 {"from": "Fuel", "to": "Combustion", "text": "generate"},
28 {"from": "Combustion", "to": "Chemical reaction", "text": "is"},
29 {"from": "Combustion", "to": "Energy", "text": "releases"},

```

```

30 {"from": "Gasoline", "to": "Fuel", "text": "is"},
31 {"from": "Car", "to": "Combustion Engine", "text": "has"},
32 {"from": "Combustion Engine", "to": "Four Steps", "text": "works"},
33 {"from": "Four Steps", "to": "Admission", "text": "consist of"},
34 {"from": "Four Steps", "to": "Compression", "text": "consist of"},
35 {"from": "Four Steps", "to": "Explosion", "text": "consist of"},
36 {"from": "Four Steps", "to": "Escape", "text": "consist of"},
37 {"from": "Battery", "to": "Spark Plug", "text": "feeds"},
38 {"from": "Combustion", "to": "Exothermic reaction", "text": "is"},
39 {"from": "Combustion", "to": "Heat", "text": "releases"},
40 {"from": "Car", "to": "Radiator", "text": "has"},
41 {"from": "Car", "to": "Battery", "text": "has"},
42 {"from": "Water", "to": "Combustion Engine", "text": "to cool"},
43 {"from": "Combustion Engine", "to": "Explosion Engine", "text": "is"},
44 {"from": "Cylinder", "to": "Piston", "text": "has"},
45 {"from": "Spark", "to": "Combustion", "text": "generate"},
46 {"from": "Radiator", "to": "Water", "text": "uses"},
47 {"from": "Energy", "to": "Piston", "text": "pushes"},
48 {"from": "Crankshaft", "to": "Wheels", "text": "moves"},
49 {"from": "Piston", "to": "Crankshaft", "text": "moves"},
50 {"from": "Gasoline", "to": "Higher Quality", "text": "needs"},
51 {"from": "Higher Quality", "to": "Pressure", "text": "supports"},
52 {"from": "Pressure", "to": "Combustion Engine", "text": "contained in the"}
53 ...
54 ]]

```

Para inserir o código json do mapa conceitual, basta acessar a aba “Body”, selecionar o campo “raw” e onde está o campo “Text”, alterar para “JSON (application/json)”. Feito isso, é só adicionar o código referente ao mapa conceitual logo abaixo. Desse modo, temos:

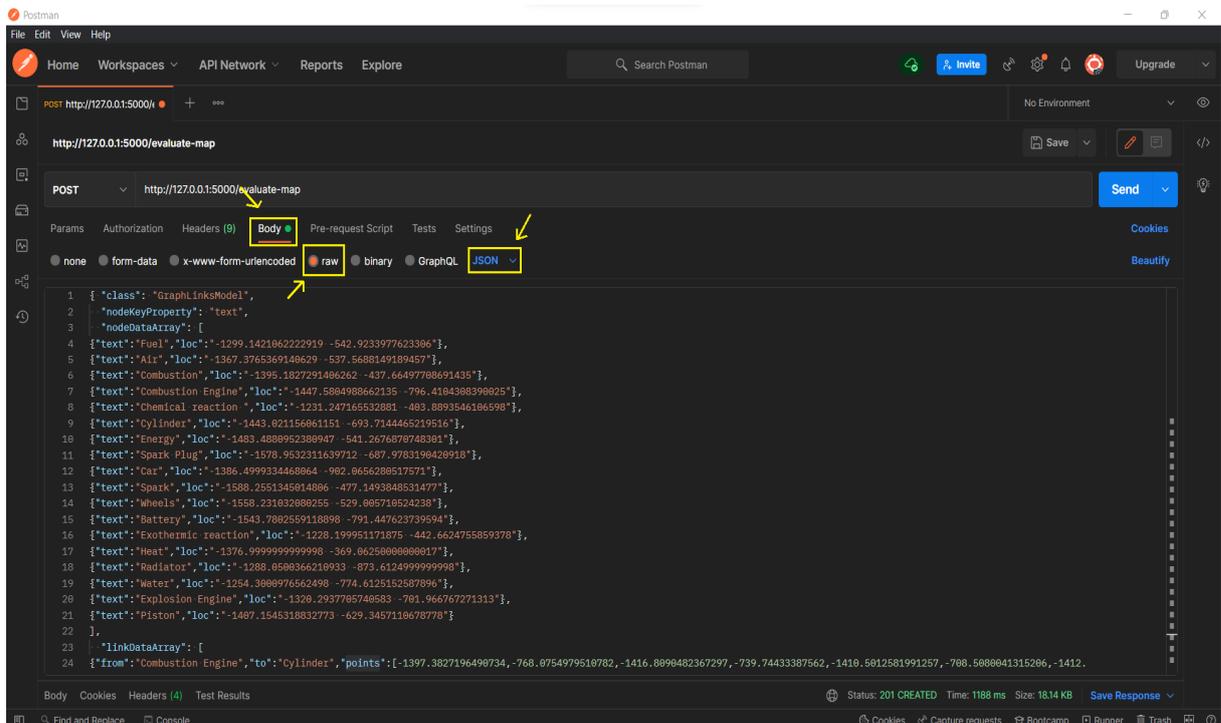
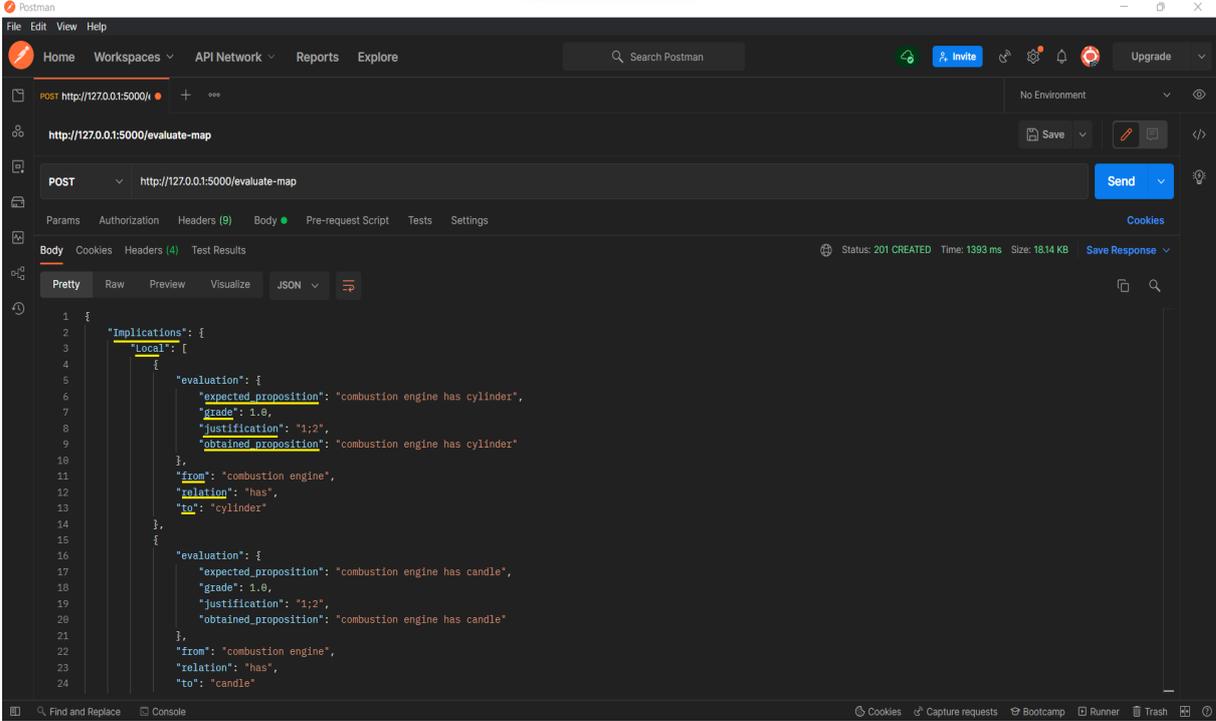


Figura 15 – Código JSON para ser requisitado ao serviço.  
Fonte: Elaborado pelo autor.

O último passo é enviar a requisição ao serviço. Para isto, basta apertar no botão “Send”. Após alguns segundos, o serviço retornará o código JSON com as implicações significantes. Este se encontra ao fim do código enviado ao serviço, como pode ser visto na figura abaixo:



```
1 {
2   "implications": {
3     "local": [
4       {
5         "evaluation": {
6           "expected_proposition": "combustion engine has cylinder",
7           "grade": 1.8,
8           "justification": "1:2",
9           "obtained_proposition": "combustion engine has cylinder"
10        },
11        "from": "combustion engine",
12        "relation": "has",
13        "to": "cylinder"
14      },
15      {
16        "evaluation": {
17          "expected_proposition": "combustion engine has candle",
18          "grade": 1.8,
19          "justification": "1:2",
20          "obtained_proposition": "combustion engine has candle"
21        },
22        "from": "combustion engine",
23        "relation": "has",
24        "to": "candle"
25      }
26    ]
27   }
28 }
```

Figura 16 – Código JSON retornado pelo serviço com as implicações significantes e demais informações.

Fonte: Elaborado pelo autor.

Na sequência, abordaremos sobre o SpaCy e Bert, bibliotecas utilizadas na API para a realização das tarefas de processamento de linguagem natural.

## 5.2 SpaCy

O spaCy ([HONNIBAL; MONTANI, 2017](#)) é uma biblioteca desenvolvida para a linguagem de programação Python, que faz processamento de linguagem natural. Sua utilização é especificamente para uso industrial, com o intuito de auxiliar a criação de aplicações que consigam processar e interpretar grandes volumes de texto, podendo ainda ser usada na extração de informações e/ou entendimento de linguagem natural ou pré-processar textos para *deep learning*. Além disso, o spaCy contém diversos modelos e uma arquitetura profissional para trabalhar com várias tarefas de processamento de linguagem natural.

Essa biblioteca tem suporte para cerca de 60 idiomas, inclusive para o português do Brasil. Porém, em nossos testes, percebemos que o modelo para o idioma português possui

um *corpus* menor do que para o inglês, ou seja, o treinamento do modelo é feito com menos textos e exemplos. Isso implica em algumas falhas na identificação de um pequeno número de relações sintáticas entre as palavras de uma frase tais como: objeto direto e indireto, verbos transitivos direto, predicativo do sujeito, entre outros. Tais classificações sintáticas são importantes para o funcionamento do nosso modelo conceitual, e, por consequência, para a análise semântica. Essas relações sintáticas entre palavras por si só já carregam uma significação. Portanto, para uma melhor amostragem da eficiência e eficácia do nosso modelo proposto, efetuamos a prova de conceito utilizando o modelo da língua inglesa. A seguir listamos as funcionalidades do processamento de linguagem natural utilizados através do spaCy no desenvolvimento do nosso modelo proposto.

### 5.2.1 Tokenização

A tokenização consiste em dividir um texto em pequenos segmentos, chamados tokens. A tokenização pode se dar no nível do documento para produzir tokens de frases ou tokenização de frase para produzir tokens de palavras ou ainda tokenização de palavras para produzir tokens de caracteres.

A tokenização geralmente é a primeira etapa do *pipeline* de pré-processamento de texto e funciona como entrada para operações de PLN subsequentes como lematização, mineração e classificação de texto, entre outros. O spaCy faz a tokenização de forma inteligente, identificando internamente se um “.” é uma pontuação, separando-o em outro token, ou se ele é parte integrante de uma abreviatura como “E.U.A.”, nesse caso mantendo-o junto à palavra. O spaCy aplica regras específicas para cada tipo de idioma. Um exemplo de texto tokenizado lexicalmente seria:

Esta é uma frase.

[‘Esta’, ‘é’, ‘uma’, ‘frase’, ‘.’]

A etapa de tokenização desempenha um papel importante na arquitetura proposta, pois permite que cada palavra seja analisada separadamente. É a primeira etapa que realizamos o processamento de linguagem natural. Com ela, podemos identificar a função sintática da palavra no contexto em que ela está inserida. A partir das funções sintáticas desempenhadas por todas as palavras contidas em uma proposição, somos capazes de classificá-la como uma implicação local ou sistêmica, por exemplo.

### 5.2.2 Lematização

A lematização é o processo que consiste em reduzir as formas flexionadas de uma palavra e, ao mesmo tempo garantir que essa forma reduzida pertença ao idioma. A forma reduzida ou palavra raiz é chamada de lema. Para o caso dos verbos, o lema é o infinitivo.

Por exemplo, “produz”, “produzido” e “produzindo” são todas formas de “produzir”. Aqui, “produzir” é o lema. Essa inflexão de uma palavra permite expressar diferentes categorias gramaticais, como tempo verbal (“produzido” x “produzir”), número (“materiais” x “material”) e assim por diante.

A lematização é necessária porque ajuda a reduzir as formas flexionadas de uma palavra para que possam ser analisadas como um único item e ainda pode auxiliar na normatização do texto. No nosso caso, a lematização também ajuda a evitar palavras duplicadas com significados semelhantes, além de facilitar a comparação para casamento de palavras durante a etapa de identificação das implicações significantes.

### 5.2.3 *Part of Speech Tagging*

O *Part of Speech* (POS) é uma função gramatical que explica como uma palavra específica é usada em uma frase. Existem ao todo oito classes gramaticais:

1. Substantivo
2. Pronome
3. Adjetivo
4. Verbo
5. Advérbio
6. Preposição
7. Conjunção
8. Interjeição

A marcação do *Part of Speech* é o processo de atribuir uma *tag* para cada *token*, dependendo do contexto da frase. As tags do POS são úteis para identificar categorias sintáticas, como substantivos ou verbos, para cada palavra. Por exemplo, vamos analisar a árvore de dependência da frase "Nina is learning piano".

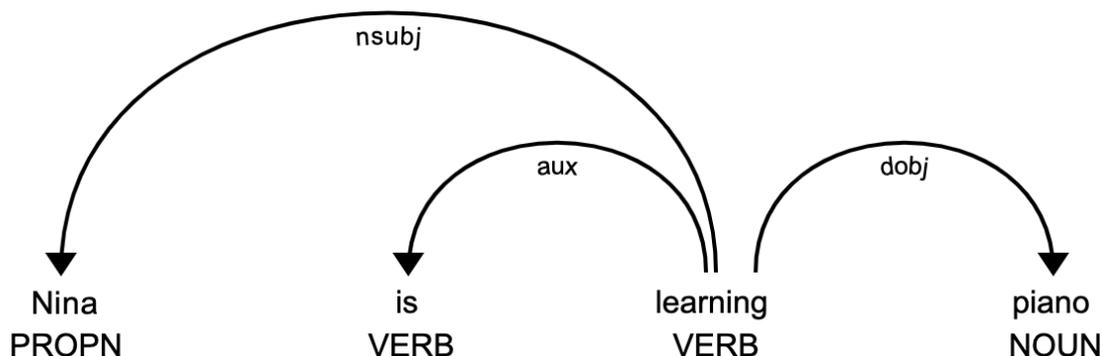


Figura 17 – Árvore de dependência para a frase do exemplo.

Fonte: Elaborado pelo autor.

Como podemos observar, a biblioteca de processamento de linguagem natural avançada é capaz de efetuar as análises morfológica e sintática. O nome próprio “Nina” é o sujeito da sentença. Ainda há a presença de dois verbos: “is” e “learning”, o primeiro tem a função de verbo auxiliar e o segundo de verbo principal. Já o nome próprio “piano” tem a função de objeto direto na análise sintática da sentença. Portanto, temos que o sujeito da frase é o nome próprio “Nina” e que ela tem uma relação de aprendizagem com o piano.

A aplicação do Part of Speech aos tokens contidos em uma proposição permite-nos identificar a classe morfológica de cada palavra. A partir dessa análise, a arquitetura pode detectar se existem verbos nas relações e substantivos nos conceitos. Após essa verificação inicial, a arquitetura pode prosseguir para a análise semântica a fim de identificar as implicações significantes.

#### 5.2.4 Vetores de Palavras e Similaridade Semântica

A similaridade é determinada através da comparação de vetores de palavras, que são representações multidimensionais do significado de uma palavra. Os vetores de palavras podem ser gerados usando um algoritmo como word2vec.

Palavras como “cachorro”, “gato” e “maça” são muito utilizadas em inglês, então elas fazem parte do vocabulário do *pipeline* do spaCy e são representadas, cada uma, por um vetor. Já a palavra “fskfsds” é desconhecida e fora do vocabulário comum do idioma - então sua representação vetorial consiste em 300 dimensões de 0, o que significa que ela é praticamente inexistente. Como nessa pesquisa precisamos de uma grande quantidade de vetores de palavras do vocabulário inglês e português para uma melhor precisão do nosso modelo, utilizamos o maior pacote de *pipeline* disponível, que inclui 685.000 vetores exclusivos.

O spaCy é qualificado para comparar dois objetos e fazer uma previsão de quão semelhantes eles são. Prever similaridade também é útil para construir sistemas de reco-

mendação ou sinalizar repetições. Por exemplo, pode-se sugerir um conteúdo de usuário semelhante ao que ele está vendo no momento ou rotular um tíquete de suporte como repetido se for muito semelhante a um já existente.

Cada token vem com um método chamado `.similarity` que permite compará-lo com outro token e determinar a semelhança entre os dois. Quanto mais próximo de 1, mais semelhantes os tokens são. É importante frisar que não há definição objetiva de similaridade. Por exemplo, as sentenças “Gosto de hambúrgueres” e “Gosto de massa” serão semelhantes dependendo do objetivo a ser alcançado. Ambos falam sobre preferências alimentares, o que os torna muito semelhantes, porém se estivermos analisando menções a alimentos, as frases serão bem diferentes, porque falam sobre alimentos muito diferentes.

O cálculo da similaridade é feito avaliando o valor do cosseno do ângulo compreendido entre dois vetores, que representam cada uma das palavras a serem analisadas. A fórmula do cálculo de similaridade é demonstrada abaixo:

$$\cos(\mathbf{A}, \mathbf{B}) = \frac{\mathbf{A}\mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \sqrt{\sum_{i=1}^n (B_i)^2}} \quad (5.1)$$

No início dos testes, começamos aceitando uma similaridade de 80% entre duas palavras e parecia que havia um perito humano lendo todas as palavras e agrupando-as. No entanto, muitas palavras que tinham contexto similares, não estavam sendo agrupadas. Por isso, fomos regredindo de 5% em 5% até chegar um ponto que já estava vindo muitas palavras sem sentido e encontramos o valor ideal foi 60%. Com 50% aparecia um grupo grande de palavras que eram similares, mas ainda havia muitos resultados indesejados. Com 55% já houve uma diminuição, mas só em 60% identificamos como sendo um valor seguro a se considerar, pois tinha uma grande quantidade de palavras por grupos e todas de alta qualidade. Com 80% os resultados possuíam uma qualidade incrível, mas poucas palavras similares eram identificadas.

### 5.3 BERT

O BERT (DEVLIN et al., 2018) é um acrônimo para Bidirecional Encoder Representations from Transformers (representações de codificador bidirecional de transformadores).

Trata-se de um algoritmo inteligente integrado pelo Google em seu mecanismo de busca no final de 2019, com o objetivo de melhorar ainda mais os resultados de busca e a experiência do usuário nas pesquisas realizadas. Ele utiliza inteligência artificial, machine learning (aprendizado de máquina) e processamento de linguagem natural (PLN). Com isso, é capaz de entender muito melhor o significado e as relações entre palavras e frases e de reconhecer o contexto completo de cada palavra utilizada nas buscas, identificando e levando em conta tanto os termos que vêm antes quanto os que vêm depois de cada

vocábulo. Por isso o termo “bidirecional” está presente em seu nome. Já os modelos de PLN tradicionais até então treinavam os modelos de linguagem na ordem da sequência de palavras (da direita para a esquerda ou da esquerda para a direita).

Essas características fazem com que o BERT seja capaz de compreender muito mais profundamente a linguagem humana, com suas gírias, variações, sinônimos, expressões, abreviações, conjunções, preposições e até mesmo erros. Sim, o algoritmo também é capaz de considerar e identificar as palavras mesmo quando escritas com a grafia errada. Um outro diferencial importante em relação ao BERT é que ele foi pré-treinado em um corpus enorme de texto que ultrapassa 33 milhões de itens.

O verdadeiro significado do que palavras comunicam só é possível por meio de uma análise profunda da rede neural utilizada pelo BERT. Por exemplo, seria muito complexo para uma máquina diferenciar a palavra “banco”, como sendo uma agência bancária, da palavra “banco”, como sendo o local para sentar em uma praça, por exemplo. O modelo contextual funciona mapeando uma representação distinta de toda a frase para entender melhor seus contextos.

Para o desenvolvimento do arcabouço, nós utilizamos o modelo de previsão de próxima palavra. O modelo consiste em substituir ou ocultar a palavra que desejamos obter pelo token [MASK]. O modelo então tenta prever o valor original das palavras mascaradas, com base no contexto fornecido pelas outras palavras não mascaradas na sequência. Com isso conseguimos ter uma ideia se a palavra utilizada pelo aluno no conceito resultante é viável de ser utilizada com base no contexto fornecido pelo conceito e relação anteriores.

Em outras palavras, seria como responder aquelas perguntas que muitos de nós recebíamos na escola, em que, dada uma sentença, precisávamos preencher as lacunas. Por exemplo, considere a sentença abaixo:

*“As \_\_\_\_\_ caem das árvores no outono.”*

Nós sabemos a resposta porque consideramos o contexto da frase. Olhamos para as palavras “caem” e “árvores” e identificamos que a palavra que falta é algo que cai das árvores. Muitas coisas caem de árvores, como frutas, galhos e folhas. Mas temos outra condição – no outono – que restringe nossa solução, de forma que a coisa mais provável de cair de uma árvore no outono são as folhas.

Como humanos, usamos uma mistura de conhecimento geral do mundo e compreensão linguística para chegar a essa conclusão. Para o BERT, esse palpite virá de “ler” muito e aprender padrões linguísticos extraordinariamente bem.

O BERT pode não saber o que são outono, árvores e folhas. Mas sabe que, dados os padrões linguísticos e o contexto dessas palavras, a resposta provavelmente será “folhas”. O

resultado desse processo, para o BERT, é uma melhor compreensão do estilo de linguagem que está sendo usado.

Nesta pesquisa, o BERT foi utilizado para dar apoio ao idioma inglês. Para o idioma português brasileiro, pesquisadores desenvolveram um modelo denominado de BERTimbau (SOUZA; NOGUEIRA; LOTUFO, 2020). Este é um modelo BERT pré-treinado que alcança performances de última geração em três tarefas de PLN: Reconhecimento de Entidade Nomeada, Semelhança Textual de Sentença e Reconhecimento de Enlace Textual. As variantes dos modelos disponíveis BERT-Base e BERT-Large Cased foram treinadas no BrWaC (Brazilian Web as Corpus), um grande corpus português, por 1.000.000 de passos, usando máscara de palavras inteiras.

## 5.4 Considerações sobre o capítulo

Neste capítulo, apresentamos de pouco mais detalhada as bibliotecas utilizadas no desenvolvimento do arcabouço, bem como os módulos respectivos aplicados em nossa solução.

O spaCy é uma biblioteca poderosa e avançada que está ganhando enorme popularidade para aplicativos de PLN devido à sua velocidade, facilidade de uso, precisão e extensibilidade.

Já a biblioteca BERT apresentou os resultados mais revolucionários na área de PLN nos últimos 5 anos, e nos permite ter uma melhor compreensão das palavras e frases dentro do contexto.

## CAPÍTULO 6

## PROVA DE CONCEITO

*Este capítulo apresentamos um experimento realizado para avaliar o modelo conceitual apresentado no capítulo 4. Esclarecemos o contexto, pré-condições necessárias e relatos das atividades realizadas pelos alunos durante o experimento, assim como alguns resultados decorrentes dessa experimentação.*

*Este capítulo está organizado da seguinte forma: a Seção*

## 6.1 Visão Geral

A fim de efetuar uma análise quantitativa e qualitativa, realizamos uma prova de conceito do MAPImp 2.0 no contexto educacional de uma turma com 24 alunos da disciplina de Organização e Representação do Conhecimento do curso de Arquivologia da Universidade Federal do Espírito Santo – UFES, sob a tutela da Prof<sup>a</sup>. Dr<sup>a</sup>. Tânia Barbosa Salles Gava.

Esta disciplina tem como objetivo capacitar o aluno na construção, organização e representação do seu próprio conhecimento, usando como apoio uma ferramenta tecnológica para construção de mapas conceituais.

A turma já havia sido introduzida ao assunto de mapas conceituais em disciplinas anteriores e, portanto, já estava familiarizada com seus elementos. A partir desse pré-conhecimento, apenas o arcabouço e seus recursos foram apresentados à turma durante o experimento.

## 6.2 Experimento

Foram realizados dois experimentos, identificados por EXP-BEGIN (realizado no primeiro dia de aula da disciplina) e EXP-END (realizado no último dia de aula da disciplina). Cada um dos experimentos foi conduzido pelo professor no ambiente de sala de aula a fim de que os alunos elaborassem um mapa conceitual sobre o assunto de Preservação Digital a partir de seu próprio conhecimento, sem o auxílio de qualquer material.

A mesma atividade foi proposta aos alunos no primeiro dia de aula e no último dia de aula do semestre letivo. A seguir, apresentamos o comando da atividade apresentado pelo professor para a atividade.

### Descrição da Atividade

Com base unicamente em seu conhecimento tácito, sem consultar nenhum material, cada aluno deve desenvolver um mapa conceitual com os conceitos abaixo. O mapa deve ser enviado pela AVA impreterivelmente até o final da aula.

- Documento Digital
- Preservação Digital
- Obsolescência Tecnológica
- Documento Arquivístico Digital
- Autenticidade
- Confiabilidade
- Hardware
- Software
- Estratégias de Preservação Digital
- Repositório Digital Confiável
- Políticas de Preservação Digital
- Tecnologias da Informação
- Acesso a longo prazo
- Fonte de Pesquisa
- Acervos Digitais

## 6.3 Premissas do experimento

Uma premissa necessária para o começo do experimento é o planejamento da atividade pelo professor. Ele tem como objeto principal a construção do mapa de referência do assunto a ser tratado com os alunos. A partir dele, são extraídos conceitos, relações entre os conceitos e implicações significantes que deverão estar contidos nos mapas construídos pelos alunos.

A Figura 18 exibe o mapa de referência elaborado durante a preparação de aula, apresentando os conceitos e relações que se espera serem construídos pelos alunos.

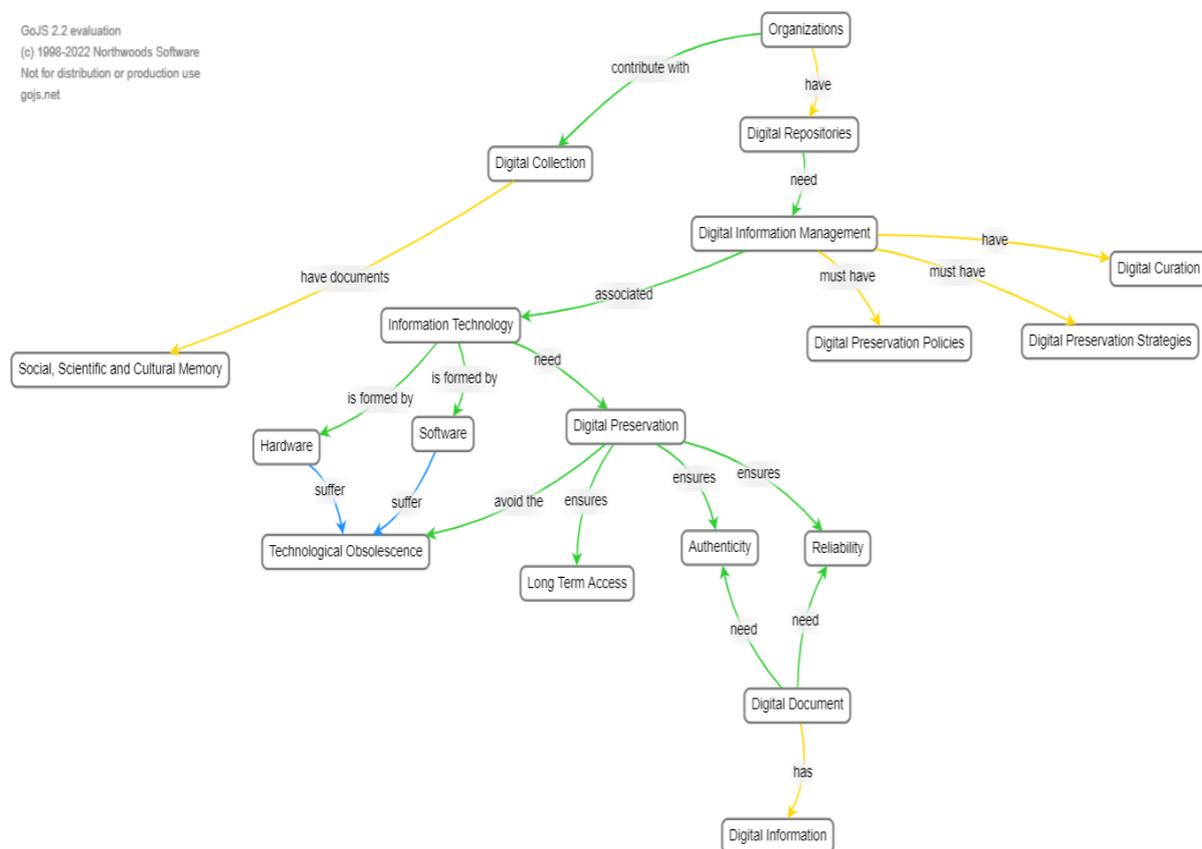


Figura 18 – Mapa de referência da atividade sobre Preservação Digital.

Fonte: Elaborado pelo autor.

Após o prazo para conclusão da atividade, todos os mapas conceituais foram enviados e salvos em um repositório.

Todos os mapas conceituais foram traduzidos manualmente para o idioma inglês, pois por se tratar de uma língua universal, seus modelos pré-treinados de processamento de linguagem natural são mais eficazes, conseguindo diagnósticos mais precisos. Como o objetivo do experimento é realizar uma prova de conceito do modelo conceitual, buscou-se evitar eventuais distorções que poderiam ser causadas por um modelo treinado por uma quantidade menor de textos, como o português.

Os passos anteriores foram executados de forma minuciosa tanto para o EXP-BEGIN quanto para o EXP-END.

## 6.4 Verificação dos Resultados

Para a **análise quantitativa**, o objetivo do experimento é avaliar a precisão da ferramenta em identificar as implicações significantes no mapa conceitual. Para isso, os mapas conceituais dos experimentos EXP-BEGIN e EXP-END foram reunidos (52 mapas) e analisados de forma manual (professor) e automática (MAPimp 2.0) segundo a presença

de implicações significantes, conforme apresentado na Seção 2.2.

A Tabela 5 apresenta os resultados da análise manual e automática para os experimentos EXP-BEGIN e EXP-END, respectivamente abreviados para BEG e END.

Aluno	Proposições		Análise do Professor						Análise da Ferramenta						Nota		% de Acerto	
			Implicações Locais		Implicações Sistêmicas		Implicações Estruturais		Implicações Locais		Implicações Sistêmicas		Implicações Estruturais					
	BEG	END	BEG	END	BEG	END	BEG	END	BEG	END	BEG	END	BEG	END	BEG	END	BEG	END
A1	18	14	5	3	11	10	0	0	5	3	11	10	0	0	7,8	7,1	100%	100%
A2	18	15	1	7	15	6	2	2	0	7	16	6	2	2	9	8,3	94%	100%
A3	30	20	0	0	26	16	4	4	0	0	27	16	2	4	10	9,9	87%	100%
A4	16	15	1	3	8	6	0	0	1	3	8	6	0	0	3,8	3,7	100%	100%
A5	20	16	5	2	9	6	2	5	5	2	9	6	2	5	7,6	8	100%	100%
A6	17	17	3	4	14	5	0	0	3	4	14	5	0	0	8,5	3,6	100%	100%
A7	23	15	0	6	21	5	2	4	0	6	21	5	2	4	10	8,3	100%	100%
A8	19	17	0	1	17	12	2	4	0	1	17	12	2	4	9,5	9,4	100%	100%
A9	16	14	0	2	15	10	0	2	0	2	15	10	0	2	7,4	7,8	100%	100%
A10	21	18	8	6	13	12	0	0	8	6	13	12	0	0	10	10	100%	100%
A11	16	13	3	4	8	9	4	0	3	4	8	9	4	0	8	7,2	100%	100%
A12	26	26	0	5	12	3	2	12	0	5	12	3	2	12	4	10	100%	100%
A13	10	15	1	4	9	9	0	0	1	4	9	9	0	0	5	8,3	100%	100%
A14	17	22	7	7	10	3	0	6	7	7	9	3	0	6	8	8,2	94%	100%
A15	20	16	2	4	16	12	2	0	2	4	16	12	2	0	8,5	8,9	100%	100%
A16	17	17	7	6	10	9	0	2	7	6	10	7	0	4	8,5	9,4	100%	88%
A17	17	14	4	4	11	10	2	0	2	7	13	7	2	0	7,9	7,8	88%	79%
A18	22	15	0	0	4	15	12	0	0	0	4	13	16	2	2,6	8,3	75%	87%
A19	17	16	8	5	9	11	0	0	8	5	7	3	2	8	7,9	8,9	88%	50%
A20	21	18	5	1	8	13	8	4	5	1	7	13	8	4	10	10	95%	100%
A21	18	15	9	4	7	8	2	2	9	4	7	8	2	2	9	7,8	100%	100%
A22	13	17	2	3	10	10	0	4	2	3	10	10	0	4	5,4	9,4	100%	100%
A23	19	21	1	3	18	10	0	7	1	3	16	8	2	7	9,5	10	89%	90%
A24	23	14	5	3	12	11	2	0	5	3	12	11	2	0	9,1	7,8	100%	100%
<b>TOTAL</b>	<b>454</b>	<b>400</b>	<b>77</b>	<b>87</b>	<b>293</b>	<b>221</b>	<b>46</b>	<b>58</b>	<b>74</b>	<b>90</b>	<b>291</b>	<b>204</b>	<b>52</b>	<b>70</b>	<b>7,79</b>	<b>8,25</b>	<b>96%</b>	<b>96%</b>

Tabela 5 – Análise dos Mapas para os experimentos EXP-BEGIN e EXP-END.

A partir dos resultados apurados, podemos observar que o MAPimp 2.0 foi capaz de identificar os três tipos de implicações significantes (local, sistêmica e estrutural) em todos os mapas conceituais. Além disso, conforme apresentado na Tabela 6, verificamos que MAPimp 2.0 atingiu em média 96% de eficácia, em ambos os experimentos, na identificação das implicações significantes. Ou seja, quase todas as implicações identificadas manualmente pelo professor, também foram identificadas pela ferramenta. Essa eficiência é calculada comparando o total de implicações identificadas pelo professor com o total identificado pela ferramenta.

Experimento	Proposições	Análise do Professor			Análise MAPimp 2.0			Nota Média	% de Acerto da Ferramenta
		Implicações Locais	Implicações Sistêmicas	Implicações Estruturais	Implicações Locais	Implicações Sistêmicas	Implicações Estruturais		
EXP-BEGIN	454	77	293	46	74	291	52	7,79	96%
EXP-END	400	87	221	58	90	204	70	8,25	96%

Tabela 6 – Análise sumarizada dos experimentos

Para a **análise qualitativa**, o objetivo é examinar se por meio das implicações significantes é possível identificar mudanças e evoluções no aprendizado dos alunos. Para tal, utilizamos os mesmos resultados do experimento anterior, porém, sobre uma perspectiva diferente.

Analisando as duas atividades desenvolvidas pela turma, no início e no final do semestre, percebemos que a quantidade total de proposições presentes nos mapas dos estudantes tiveram uma queda de 12%: um total de 454 proposições em EXP-BEGIN e de 400 proposições em EXP-END. No entanto, percebemos que mesmo com um menor número de proposições representadas no EXP-END, a média das notas da turma ainda teve um pequeno aumento. Isso pode ser justificado pelo fato de que em EXP-BEGIN os relacionamentos do mapa representaram um conhecimento desorganizado sobre o assunto e, por isso, as relações não eram tão claras e significativas. Exemplo disso são proposições formadas por textos extensos na tentativa de representar esse conhecimento não estruturado. Desse modo, em EXP-END, os alunos organizaram e estruturaram melhor esse conhecimento a partir do conteúdo explorado durante o semestre, representando a mudança em relação ao conhecimento inicial.

Partindo para a análise específica das notas atribuídas pela ferramenta aos mapas do EXP-BEGIN e EXP-END, percebemos que 13 alunos apresentaram evoluções e conseguiram uma nota maior em EXP-END ou manter a nota máxima, o que representa mais da metade da turma de 24 alunos. Além disso, 7 alunos sustentaram a média da nota da primeira atividade na atividade final, considerando a margem de até 1,0 ponto para menos. 3 alunos apresentaram uma redução entre 1 e 2 pontos, e apenas 1 aluno apresentou uma redução significativa.

Segundo os experimentos realizados, notamos que a turma adquiriu um certo conhecimento sobre o assunto abordado no início do semestre, e a ferramenta teve potencial para identificar essa mudança, sinalizando uma mudança no conhecimento. Podemos refletir que a construção do conhecimento proposta por Piaget está perceptível em todos os mapas, porém é percebida de maneira mais clara onde o aprendizado foi construído ao passar dos meses, onde novos conhecimentos são assimilados, acomodados e adaptados. Vale ressaltar que o experimento demonstrou a construção de implicações locais a partir do conhecimento inicialmente desestruturado e um aumento de implicações estruturais a medida em que o conhecimento é assimilado, mas ainda não há evidências sobre mudanças no nível de implicações no decorrer do aprendizado.

É importante salientar que esses experimentos partiram de uma dinâmica em que os conceitos a serem utilizados foram pré-determinados pela descrição da atividade. Por outro lado, se a dinâmica permitir que os alunos enumerem conceitos novos, pode ser que apareçam termos não previstos pelo professor ou que destoem daquilo que foi proposto.

## 6.5 Considerações sobre o capítulo

Neste capítulo apresentamos a experimentação em um ambiente de sala de aula com o propósito de validar o modelo conceitual proposto. Ao término dos experimentos

foi possível verificar as qualidades, melhoradas a serem implementadas, bem como refletir criticamente acerca dos resultados e extrair conclusões que serão tratadas no próximo capítulo.

## CAPÍTULO 7

## CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

*Este capítulo apresenta algumas considerações de todo o trabalho desenvolvido por esta pesquisa, bem como a discussão para trabalhos futuros essenciais para a continuidade e aprimoramento desta pesquisa.*

*Este capítulo está organizado da seguinte forma: a Seção 7.1 apresenta as considerações finais; e a Seção 7.2 discutem trabalhos futuros.*

## 7.1 Considerações Finais

Para responder ou validar as hipóteses da pesquisa apresentadas no Capítulo 1, esta pesquisa iniciou-se com uma revisão bibliográfica sobre as abordagens tecnológicas direcionadas à utilização de análise semântica em mapas conceituais. A revisão da literatura resultou em uma categorização para melhor identificar e analisar as funcionalidades e características das abordagens tecnológicas neste contexto.

A categorização obtida permitiu uma visualização e análise de forma abrangente e precisa das principais características adotadas em cada abordagem, e ainda auxiliou na definição do nosso modelo conceitual. A partir do modelo conceitual, percebemos que nenhum dos trabalhos relacionados contemplou as características adotadas pelo MAPimp 2.0 após a categorização, pois combina:

- (i) Análise sintática;
- (ii) Análise semântica;
- (iii) Sistema de pontuação;
- (iv) Processo de avaliação automático.

A partir do modelo conceitual proposto, definimos uma arquitetura tecnológica para atender ao objetivo desta pesquisa. Essa arquitetura computacional foi aplicada no desenvolvimento da ferramenta MAPimp 2.0, que está disponível para ser utilizada publicamente como uma API.

Segundo os experimentos realizados com a ferramenta MAPimp 2.0 em um ambiente educacional, apresentados no Capítulo 6, podemos observar a qualidade e quantidade de informações que podem ser extraídas de mapa conceitual construído, a partir da análise das implicações significantes. Com base nos resultados obtidos, concluímos que a avaliação de

um mapa pela ferramenta MAPimp 2.0 apresenta resultados aceitáveis quando comparado com avaliações por especialistas.

Além dos pontos apresentados, ressaltamos que nenhum trabalho relacionado está disponível publicamente para uso, extensão ou serviço. Nesse contexto, a ferramenta MAPimp 2.0 oferece uma grande contribuição para a Educação e Pesquisa, uma vez que a API é composta por diversas bibliotecas de extração de informações disponíveis para uso e extensão para outros projetos.

Ainda de acordo com os experimentos realizados, podemos observar algumas características dos modelos e bibliotecas desenvolvidos. Comparado com outras bibliotecas, concluímos que MAPimp 2.0, além dos recursos trazidos em sua primeira geração, pode:

- (i) Aplicar técnicas de processamento de linguagem natural de previsão e casamento de palavras, com auxílio de um modelo treinado por redes neurais.
- (ii) Aplicar análise sintática e semântica simultaneamente;
- (iii) Adotar um método de avaliação por proposições.
- (iv) Gerar um texto a partir do conteúdo do mapa conceitual;
- (v) Extrair diversas informações do mapa conceitual, como número de aproximações, lista de implicações, justificativas para a avaliação, nota final, entre outros.

Portanto, considerando a pesquisa desenvolvida ao longo desta dissertação, podemos concluir que as seguintes hipóteses foram validadas:

- (i) Com base no Capítulo 6, concluímos que é possível criar uma ferramenta capaz de identificar implicações significantes, utilizando análise semântica, em mapas conceituais de estilo científico de forma a aprimorar a análise e a avaliação da aprendizagem. Observando os experimentos apresentados na Seção 6.4, concluímos que é possível utilizar a análise de implicações como alternativa para avaliação do aprendizado.
- (ii) Com base na arquitetura tecnológica apresentada no Capítulo 4, podemos concluir que o uso de modelos de linguagem pré-treinados para processamento de linguagem natural pode dar uma dimensão semântica na avaliação de mapas para aplicação multilíngue nos idiomas português e inglês.
- (iii) Com base nos experimentos apresentados na Seção 6.4, a comparação de similaridade de palavras entre os mapas dos alunos de uma mesma turma pode evidenciar um fechamento contextual, de forma a ampliar a dimensão semântica.

Por fim, devemos considerar que os trabalhos dedicados à análise semântica em mapas conceituais são relativamente novos e ainda estão em evolução. O MAPimp 2.0 tem mostrado resultados promissores, embora existam muitos desafios a serem superados.

## 7.2 Trabalhos Futuros

Existem muitos desafios derivados deste trabalho, de modo a contribuir para seu aperfeiçoamento e precisão, além de possibilitar uma melhor experiência do aluno no processo de aprendizagem e colaborar ainda mais para a Educação e Pesquisa.

Trabalhos futuros serão focados na qualidade da ferramenta MAPimp 2.0, no desenvolvimento de bibliotecas para a API e nos estudos de análise semântica de mapas conceituais.

Precisamos dedicar nossos esforços para garantir a qualidade das funcionalidades desenvolvidas. Por isso, nossos trabalhos futuros estarão relacionados a testes e melhorias das funcionalidades, disponibilização de novos serviços, usabilidade e eficiência da ferramenta MAPimp 2.0. Além disso pretendemos ampliar a análise da ferramenta para outros idiomas.

Destacamos a importância da análise semântica no processo de avaliação de mapas conceituais, por meio da utilização das implicações significantes de Piaget. No entanto, além desta dissertação, estudos mais precisos e aprofundados sobre o tema precisam ser realizados, podendo trazer ainda mais contribuições para a educação em geral.

Por fim, para resultados mais amplos e precisos, queremos disponibilizar a ferramenta para professores e alunos da rede pública estadual.

## REFERÊNCIAS

- AGUIAR, C. Z. de; CURY, D.; ZOUAQA, A. Towards technological approaches for concept maps mining from text. *CLEI Electronic Journal*, v. 21, n. 1, 2018. Citado 2 vezes nas páginas 20 e 27.
- AGUIAR, Z. C.; CURY, D. A categorization of technological approaches to concept maps construction. In: *2016 XI Latin American Conference on Learning Objects and Technology (LACLO)*. IEEE, 2016. p. 1–9. ISBN 978-1-5090-6149-5. Disponível em: <<http://ieeexplore.ieee.org/document/7751743/>>. Citado na página 22.
- ANDER-EGG, E. *Introducción a las técnicas de investigación social para trabajadores sociales*. [S.l.], 1978. Citado na página 22.
- AUSUBEL, D. P. *Educational Psychology, A Cognitive View*. [S.l.]: Holt, Rinehart and Winston, 1978. Citado 2 vezes nas páginas 18 e 27.
- BOGUSKI, R. R.; CURY, D. Usando regras de associação para a identificação de falhas conceituais. In: *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*. [S.l.: s.n.], 2018. v. 29, n. 1, p. 1443. Citado na página 44.
- BOGUSKI, R. R.; CURY, D.; GAVA, T. Tom: An intelligent tutor for the construction of knowledge represented in concept maps. In: IEEE. *2019 IEEE Frontiers in Education Conference (FIE)*. [S.l.], 2019. p. 1–7. Citado 3 vezes nas páginas 26, 38 e 54.
- BOSS, S.; MATOS, E.; ANDRADE, A. Soluções computacionais para avaliação da aprendizagem por mapas conceituais: um mapeamento sistemático. In: *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*. [S.l.: s.n.], 2019. v. 30, n. 1, p. 399. Citado na página 19.
- DEVLIN, J. et al. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018. Citado 2 vezes nas páginas 43 e 67.
- DUTRA, Í. M.; FAGUNDES, L. d. C.; CAÑAS, A. Un enfoque constructivista para el uso de mapas conceptuales en educación a distancia de profesores. In: *Concept Maps: Theory, Methodology, Technology. Proceedings of the First International Conference on Concept Mapping*. [S.l.: s.n.], 2004. v. 1, p. 217–226. Citado na página 30.
- DUTRA, Í. M. et al. Uma base de dados para compartilhamento de experiências no uso de mapas conceituais no acompanhamento de processos de conceituação. *RENOTE*, v. 4, n. 2, 2006. Citado na página 19.
- DUTRA Ítalo M. 2018. [Online; acessado em 13-Outubro-2021]. Disponível em: <[http://peadalvorada5.pbworks.com/f/italo\\_mapas.pdf](http://peadalvorada5.pbworks.com/f/italo_mapas.pdf)>. Citado 4 vezes nas páginas 12, 19, 29 e 30.

- ELSAYED, O. et al. An on-line framework for experimenting with concept maps. In: IEEE. *2019 18th International Conference on Information Technology Based Higher Education and Training (ITHET)*. [S.l.], 2019. p. 1–8. Citado 3 vezes nas páginas 38, 39 e 47.
- FUENTES-LORENZO, D. et al. Building concept maps by adapting semantic distance metrics to wikipedia. *Education for Information*, IOS Press, v. 35, n. 3, p. 209–240, 2019. Citado na página 38.
- GIL, A. C. *Métodos e técnicas de pesquisa social*. [S.l.]: 6. ed. Editora Atlas SA, 2008. Citado na página 23.
- HAO, J.-X.; YAN, A.; CHI-WAI, R. A semantic analysis method for concept map-based knowledge modeling. In: ATLANTIS PRESS. *1st International Conference on E-Business Intelligence (ICEBI 2010)*. [S.l.], 2010. p. 117–123. Citado na página 44.
- HOJAS-MAZO, W. et al. Semantic processing method to improve a query-based approach for mining concept maps. In: SPRINGER. *International Conference on Computer Science, Electronics and Industrial Engineering (CSEI)*. [S.l.], 2019. p. 22–35. Citado na página 39.
- HONNIBAL, M.; MONTANI, I. Natural language understanding with bloom embeddings, convolutional neural networks and incremental parsing. *Unpublished software application*. <https://spacy.io>, 2017. Citado 2 vezes nas páginas 53 e 63.
- LIMA, J. L. O.; ALVARES, L. Organização e representação da informação e do conhecimento. *Organização da informação e do conhecimento: conceitos, subsídios interdisciplinares e aplicações*. São Paulo: B4 Editores, v. 248, p. 21–48, 2012. Citado na página 17.
- LIMONGELLI, C. et al. A framework for comparing concept maps. In: IEEE. *2017 16th International Conference on Information Technology Based Higher Education and Training (ITHET)*. [S.l.], 2017. p. 1–6. Citado 2 vezes nas páginas 39 e 47.
- MANSKE, S.; HOPPE, H. U. Managing knowledge diversity: towards automatic semantic group formation. In: IEEE. *2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT)*. [S.l.], 2017. p. 330–332. Citado na página 38.
- MARCONI, M. de A.; LAKATOS, E. M. et al. *Metodologia científica*. [S.l.]: Atlas, 2007. Citado na página 23.
- MIKOLOV, T. et al. Distributed representations of words and phrases and their compositionality. In: *Advances in neural information processing systems*. [S.l.: s.n.], 2013. p. 3111–3119. Citado na página 44.
- MOREIRA, C. A. e. D. C. R. Ferramenta computacional multilíngue para analisar significações em mapas conceituais. *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação - SBIE)*, v. 30, n. 1, p. 419, 2019. ISSN 2316-6533. Disponível em: <<https://www.br-ie.org/pub/index.php/sbie/article/view/8746>>. Citado 4 vezes nas páginas 19, 20, 46 e 47.
- MOREIRA, M. A. Mapas conceituais e aprendizagem significativa (concept maps and meaningful learning). *Aprendizagem significativa, organizadores prévios, mapas conceituais, digramas V e Unidades de ensino potencialmente significativas*, v. 41, 2012. Citado na página 19.

- MOREIRA, R.; AGUIAR, C.; CURY, D. Ferramenta computacional multilíngue para analisar significações em mapas conceituais. In: *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*. [S.l.: s.n.], 2019. v. 30, n. 1, p. 419. Citado 3 vezes nas páginas 19, 31 e 54.
- NOVAK, J. D.; CANAS, A. J. *The Theory Underlying Concept Maps and How to Construct Them*. Institute for Human and Machine Cognition., 2006. Citado na página 27.
- NOVAK, J. D.; GOWIN, D. B.; BOB, G. D. *Learning how to learn*. [S.l.]: cambridge University press, 1984. Citado na página 18.
- PETERSEN, K. et al. Systematic mapping studies in software engineering. In: *12th International Conference on Evaluation and Assessment in Software Engineering (EASE) 12*. [S.l.: s.n.], 2008. p. 1–10. Citado na página 34.
- PIAGET, J.; GARCIA. *Hacia una Lógica de Significaciones*. [S.l.]: Gedisa, Mexico, 1989. Citado 2 vezes nas páginas 18 e 29.
- QUIVY, R.; CAMPENHOUDT, L. V. *Manual de investigação em ciências sociais*. Gradiva Lisboa, 1992. Citado na página 23.
- RIOS, C. A. e. D. C. P. Uma abordagem construtivista para a identificar o conhecimento usando mapas conceituais. *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação - SBIE)*, v. 28, n. 1, p. 394, 2017. ISSN 2316-6533. Disponível em: <<https://www.br-ie.org/pub/index.php/sbie/article/view/7568>>. Citado na página 19.
- RIOS, P.; CURY, D.; DUTRA, Í. M. Automatizando uma argumentação construtivista por meio dos mapas conceituais. In: *XX Congreso Internacional de Informática Educativa (TISE'2015)*. Disponível em: <http://www.tise.cl/volumen11/TISE2015/157-162.pdf>. [Google Scholar]. [S.l.: s.n.], 2015. Citado 4 vezes nas páginas 19, 27, 30 e 31.
- SALEM, S. B.; CHENITI-BELCADHI, L.; BRAHAM, R. A concept map based scenario for assessment of short and open answer questions. In: *IEEE. 2017 International Conference on Engineering & MIS (ICEMIS)*. [S.l.], 2017. p. 1–7. Citado 3 vezes nas páginas 38, 39 e 47.
- SAMPIERI, R.; COLLADO, C.; LUCIO, M. Coleta e análise dos dados qualitativos. *Metodologia de pesquisa. Porto Alegre: Penso*, 2013. Citado na página 22.
- SORGENTE, A.; VETTIGLI, G.; MELE, F. Automatic extraction of cause-effect relations in natural language text. *DART@ AI\* IA*, Citeseer, v. 2013, p. 37–48, 2013. Citado na página 46.
- SOUZA, F.; NOGUEIRA, R.; LOTUFO, R. BERTimbau: pretrained BERT models for Brazilian Portuguese. In: *9th Brazilian Conference on Intelligent Systems, BRACIS, Rio Grande do Sul, Brazil, October 20-23 (to appear)*. [S.l.: s.n.], 2020. Citado na página 69.
- VASQUES, D. G.; MARTINS, P. S.; REZENDE, S. O. A semantic approach to uncovering implicit relationships in textual databases. In: *IEEE. 2018 XLIV Latin American Computer Conference (CLEI)*. [S.l.], 2018. p. 490–499. Citado na página 38.