



UNIVERSIDADE FEDERAL RURAL DO SEMI-ÁRIDO
UNIVERSIDADE DO ESTADO DO RIO GRANDE DO NORTE
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA
COMPUTAÇÃO



RAIMUNDO NONATO BEZERRA NETO

**Sistema Classificador de Mapas Conceituais: Uma
Arquitetura Computacional Baseada em Processamento
de Linguagem Natural**

Mossoró-RN

2019

RAIMUNDO NONATO BEZERRA NETO

**Sistema Classificador de Mapas Conceituais: Uma
Arquitetura Computacional Baseada em Processamento
de Linguagem Natural**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação - associação ampla entre a Universidade do Estado do Rio Grande do Norte e a Universidade Federal Rural do Semi-Árido, para a obtenção do título de Mestre em Ciência da Computação.

Linha de Pesquisa: Tecnologias Aplicada à Educação e à Saúde

Orientador: Rommel Wladimir de Lima, Prof. Dr.

Mossoró-RN

2019

Resumo

Com a democratização da informação, o modelo de ensino tradicional ganha um desafio ainda maior, o de manter os alunos motivados em sala de aula e, no mesmo momento, otimizar o processo de ensino aprendizagem para os professores. Entre as inúmeras alternativas que estão sendo propostas, a Teoria da Aprendizagem Significativa (TAS) aliada ao estudo sobre Mapas Conceituais conduz a um ambiente propício para esse cenário que se anuncia. A utilização de Mapas Conceituais vem ganhando espaço no ambiente educacional, permitindo a modelagem do conhecimento através da organização de conceitos em uma estrutura proposicional. Ao passar dos anos, houve um considerável interesse da comunidade acadêmica na utilização de recursos computacionais para interagir com Mapas Conceituais, criação, comparação, análise evolutiva do aprendiz, processamento de linguagem natural e compartilhamento de mapas são algumas das tarefas mais executadas no contexto de otimização. O objetivo deste trabalho é apresentar uma arquitetura computacional focada na análise da linguagem natural, que utiliza os relacionamentos textuais produzidos pelos Mapas Conceituais e desenvolvidos na ferramenta *CmapTools* para classificar o nível de cognição do aluno e, aditivamente, expor esses resultados de forma analítica para servir de subsídio para o especialista.

Palavras-chave: Aprendizagem Significativa; Mapa Conceitual; Processamento de Linguagem Natural.

Abstract

With the democratization of information, the traditional teaching model faces an even greater challenge of keeping students motivated in the classroom while optimizing the teaching-learning process for teachers. Among the many alternatives that are being proposed, the Meaningful Learning Theory (TAS) combined with the study on Concept Maps leads to an environment conducive to this scenario that is heralded. The use of Concept Maps has been gaining ground in the educational environment, allowing the modeling of knowledge through the organization of concepts in a propositional structure. Over the years, there has been considerable interest from the academic community in using computing resources to interact with Concept Maps, creating, comparing, learner evolutionary analysis, natural language processing, and map sharing are some of the most commonly performed tasks in the optimization context. The aim of this paper is to present a computational architecture focused on natural language analysis, which uses the textual relationships produced by the Concept Maps and developed in the CmapTools tool to classify the student's cognition level and, additively, to expose these results analytically to serve allowance for the specialist.

Keywords: Significant Learning; Conceptual Map; Natural Language Processing.

Lista de ilustrações

Figura 1 – Pirâmide de Aprendizagem de William Glasse	18
Figura 2 – Elementos do processo de ensino-aprendizagem	24
Figura 3 – Ilustração do processo de aprendizagem Mecânica e de Aprendizagem significativa	25
Figura 4 – Comparação entre aprendizagem mecânica e significativa de acordo com a Teoria de Ausubel	26
Figura 5 – Uma versão reduzida da teoria de aprendizagem de Ausubel	27
Figura 6 – Evolução do mapeamento conceitual num contexto histórico	28
Figura 7 – Estrutura de Criação de um Mapa Conceitual	29
Figura 8 – Árvore Sintática	32
Figura 9 – Análise de Dependência Sintática	35
Figura 10 – Reconhecimento de Entidade Nomeada (REN)	35
Figura 11 – Quantidade de mapas classificados por tema e por nível	38
Figura 12 – Distância no <i>ranking</i> de classificação entre Novak e os Especialistas	42
Figura 13 – Distância no <i>ranking</i> de classificação entre Novak e os Não-Especialistas	43
Figura 14 – Modelo Conceitual da Fermenta de Clusterização de Mapas Conceituais	45
Figura 15 – Comparação do Professor versus Comparação Automática	46
Figura 16 – Tecnologias Emergentes em 2018	47
Figura 17 – Citações sobre o tópico “ <i>natural language processing</i> ” por ano	47
Figura 18 – Comparação de quantidades de caracteres	48
Figura 19 – Twitter com destaque para EN	49
Figura 20 – Arquitetura do Sistema	52
Figura 21 – Saída do Mapa Conceitual Base	54
Figura 22 – Mapa Conceitual Base	55
Figura 23 – Texto com identificação de abreviações	56
Figura 24 – Trecho da Base de Dados NLP classificada e dividida em frases	57
Figura 25 – Trecho da Base de Dados Ranqueada	57
Figura 26 – Modelagem dos Processos	58
Figura 27 – Arquivo de Saída “remontado”	59
Figura 28 – Arquivo de Saída “remontado” sem a palavra passe	60
Figura 29 – Arquivo de Saída do aluno Alfa	61
Figura 30 – Classificação Semântica do aluno Alfa	61
Figura 31 – Classificação Semântica errônea dos verbos do aluno Alfa	63
Figura 32 – Classificação Semântica correta dos verbos do aluno Alfa	63
Figura 33 – <i>Tokenização</i> do mapa do Aluno Alfa	64
Figura 34 – Remoção dos <i>Stop Words</i> do mapa do Aluno Alfa	65

Figura 35 – Lista dos <i>StopWords</i> do <i>corpus</i> Floresta	65
Figura 36 – Proposições do Aluno Alfa após aplicação do Algoritmo de <i>stemming</i> .	66
Figura 37 – Lista de proposições válidas do Aluno Alfa	67
Figura 38 – Lista de proposições inválidas do Aluno	67
Figura 39 – <i>Scores</i> do Aluno Alfa	68
Figura 40 – Classificação do mapa do Aluno Alfa	69

Lista de tabelas

Tabela 1 – Comparação entre texto normal e texto marcado	33
Tabela 2 – Scores do Mapa Conceitual	39
Tabela 3 – Exemplo de tabela de Clareza Proposicional	40
Tabela 4 – Condições necessárias para classificar a taxonomia topológica	44
Tabela 5 – Comparação dos Trabalhos Relacionados	50
Tabela 6 – <i>Tags</i> da Biblioteca spaCy	62

Lista de abreviaturas e siglas

AVA	Ambiente Virtual de Aprendizagem
IA	Inteligência Artificial
KDD	<i>Knowledge Discovery in Databases</i>
MC	Mapas Conceituais
NLTK	<i>Natural Language Toolkit</i>
NLP	<i>Natural Language Processing</i>
PLN	Processamento de Linguagem Natural
PET	Programa de Educação Tutorial
POS	<i>Parts of Speech</i>
RAM	<i>Random Access Memory</i>
REN	Reconhecimento de Entidades Nomeadas
ROM	<i>Read-Only Memory</i>
SCCM	<i>Scoring Criteria for Concept Maps</i>
SI	Sistemas de Informação
TAS	Teoria da Aprendizagem Significativa
TDC	Teoria da Dupla Codificação
TDIC	Tecnologias Digitais de Informação e Comunicação
UERN	Universidade do Estado do Rio Grande do Norte
UNESCO	Organização das Nações Unidas para a Educação, a Ciência e a Cultura

Sumário

1	INTRODUÇÃO	16
1.1	CONTEXTO	17
1.2	PROBLEMÁTICA	19
1.3	OBJETIVOS	20
1.3.1	Objetivo Geral	20
1.3.2	Objetivos Específicos	20
1.4	METODOLOGIA	20
1.5	MOTIVAÇÃO	21
1.6	ORGANIZAÇÃO DO DOCUMENTO	21
2	ASPECTOS CONCEITUAIS	23
2.1	TEORIAS DA APRENDIZAGEM SIGNIFICATIVA	23
2.1.1	Aprendizagem Significativa x Aprendizagem Mecânica	24
2.1.2	Diferenciação Progressiva	26
2.1.3	Reconciliação Integrativa	26
2.2	MAPAS CONCEITUAIS	27
2.3	PROCESSAMENTO DE LINGUAGEM NATURAL	31
2.4	PROGRAMAÇÃO PARA PLN	33
2.4.1	Biblioteca NLTK	33
2.4.2	Biblioteca spaCy	34
3	TRABALHOS RELACIONADOS	36
3.1	APRENDIZAGEM SIGNIFICATIVA	36
3.2	MAPAS CONCEITUAIS	39
3.3	PROCESSAMENTO DE LINGUAGEM NATURAL	46
4	DESENVOLVIMENTO DA ARQUITETURA COMPUTACIONAL DE CLASSIFICAÇÃO DE MAPAS CONCEITUAIS	51
4.1	VISÃO GERAL	51
4.2	A ARQUITETURA	52
4.3	VALIDAÇÃO	53
4.3.1	Criação Dos Mapas Conceituais	53
4.3.2	Exportação do Arquivo Cmap	54
4.3.3	Base de Dados	55
4.3.4	Geração dds Arquivos de Entrada	58
4.3.5	Análise Geral das Proposições	59

4.3.6	Normalização e Segmentação de Texto	64
4.3.7	Análise Contextual	66
4.3.8	Geração dos <i>Scores</i>	67
4.3.9	Geração do Nível da Taxonomia Topológica	68
4.4	RESULTADOS	69
5	CONSIDERAÇÕES FINAIS	70
5.1	CONCLUSÃO	70
5.2	LIMITAÇÕES	71
5.3	TRABALHOS FUTUROS	72
	REFERÊNCIAS	73
	APÊNDICES	79
	APÊNDICE A – INSTRUÇÕES	80
	APÊNDICE B – CLASSIFICAÇÕES GRAMATICAIS	81
	APÊNDICE C – AVALIAÇÕES DO ESPECIALISTA	84
	APÊNDICE D – VALIDAÇÕES DOS SCORES	88
	APÊNDICE E – CLASSIFICAÇÃO DA TAXONOMIA TOPOLO- GICA	92
	ANEXOS	94
	ANEXO A – SCORING CRITERIA FOR CONCEPT MAPS	95
	ANEXO B – TAXONOMIA TOPOLÓGICA	96

1 Introdução

O ensino superior tradicional vem passando por mudanças estruturais e metodológicas nos últimos tempos, mudanças essas que desafiam a forma de ensinar, de aprender e que atinge diretamente a sua sustentabilidade. Salas de aula cada vez mais vazias são reflexos da maneira passiva de se lecionar, nesse mesmo ambiente onde acredita-se que os alunos aprendem no mesmo ritmo e da mesma forma. Segundo Valente (2014) o modelo de universidade que faz pesquisa, gera conhecimento e distribui este conhecimento para poucos, já não se sustenta mais. Dessa forma, novas metodologias surgiram para preencher lacunas deixadas pelos modelos tradicionais, procedimentos que se utilizam de novas tecnologias proporcionam ambientes de aprendizagem mais flexíveis. Esses novos ambientes geram um ensino mais dinâmico, envolvido em um aprendizado mais ativo conforme afirma Fava (2014). Seguindo esse propósito, a Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) indica a efetiva utilização de ambientes educacionais tecnológicos, trazendo relação direta entre educação, tecnologia e bem-estar social. Continuando nesse contexto Magrini *et al.* (2010) afirma que “*para viver, aprender e trabalhar bem em uma sociedade cada vez mais complexa, rica em informação e baseada em conhecimento, alunos e professores devem usar a tecnologia de forma efetiva [...]*”.

"[...] o uso da tecnologia pode permitir que os alunos tornem-se usuários qualificados da informação; audiência crítica da informação; solucionadores de problemas e tomadores de decisão; usuários criativos e efetivos de ferramentas de produtividade; comunicadores, colaboradores, editores e produtores de informação; cidadãos informados, responsáveis e capazes de contribuir com a sociedade" conforme cita (MAGRINI *et al.*, 2010).

Apoiado nesse cenário, Moreira (2006) explica que a integração desses ambientes tecnológicos apoiado na metodologia de Mapas Conceituais leva o aluno a adquirir a chamada “aprendizagem significativa” de Ausubel *et al.* (1968). A aprendizagem é considerada significativa no momento em que uma nova informação se transforma em significado para o aprendiz, mediante conexão com aspectos relevantes da estrutura cognitiva preexistente.

Diante do exposto, Kowata (2010) afirma que o desenvolvimento do conhecimento em um ambiente de ensino e aprendizagem deve ser facilitado, utilizando mecanismos tecnológicos, e sem dissociar o domínio técnico do pedagógico.

Nesse trabalho, usamos recursos de Processamento de Linguagem Natural para processar as informações contidas em Mapas Conceituais, com o objetivo de identificar a presença de aprendizagem significativa. De forma mais específica, analisamos cada proposição identificando similaridades e contextos específicos, com técnicas que podem identificar tais característica utilizando a análise textual.

1.1 CONTEXTO

O construtivismo, movimento artístico que preconiza a integração entre as técnicas artesanais e a produção industrial, tem várias ramificações e, todas essas vertentes consentem em considerar que o aprendiz relaciona a informação recebida ao seu conhecimento empírico sobre o assunto. A construção do conhecimento é própria, sendo moldada de acordo com as experiências do indivíduo ao longo de sua vida. Segundo Ausubel, Novak e Hanesian (1980), o indivíduo consegue assimilar conceitos e/ou conteúdo de forma mais eficiente quando realiza uma abordagem mais geral sobre um determinado tema, ao invés de construir significados partindo de temas mais específicos.

Seguindo esse raciocínio, o atual modelo pedagógico utilizado na universidade moderna, onde o professor é mero transmissor de conteúdo, está se tornando obsoleto, sobre isso Tapscott e Williams (2010) declara que esse padrão baseado na transmissão pode ter sido apropriado para uma economia e uma geração anterior, onde a necessidade e a rapidez da informação seguia outros contornos. Hoje, esse modelo deixa de atender às necessidades de uma nova geração de estudantes que estão imergidos em um cenário global do conhecimento, gerando assim um elevado índice de evasão nas instituições de ensino superior no Brasil.

Para retratar esse cenário nacional, Filho *et al.* (2007) analisaram as informações do MEC no contexto de 2000 a 2005 e constataram que, a média anual de evasão escolar no ensino superior girava em torno de 22%. O mesmo estudo preocupou-se em realizar comparativo sobre a evasão a nível de América Latina e em outros países mais desenvolvidos, observou-se que as taxas de evasão brasileira eram menores do que em outros países da América Latina. Contudo, ao se comparar com países orientais como Japão e Turquia e países como França, Itália e Suécia, essas taxas eram bem mais elevadas. Nessa concepção, pressupõe que o problema da evasão é internacional e o Brasil acompanha essa escalada.

Do mesmo modo Puentes e Aquino (2008), em sua pesquisa sobre gerenciamento de tempo em sala de aula universitária, constatam que 37,31% do tempo efetivo direcionado ao processo de ensino aprendizagem são desperdiçados por alunos e professores com ausências, atrasos no início da aula, interrupções e falta de permanência na sala. Gerando uma necessidade constante de otimização desse tempo.

Igualmente, Bloom (1944) e Bloom (1972) comprovou que a capacidade humana de aprendizagem ocorre de forma diferente entre os indivíduos. Ao longo da história julgou-se que o motivo pelo qual uma porcentagem de discentes atingia uma performance melhor do que outros estava associada às situações e variáveis existentes fora do ambiente educacional o qual, nas mesmas condições de aprendizagem, todos aprenderiam com a mesma capacidade e argúcia o conteúdo.

Para ratificar essa teoria Bloom *et al.* (1971) evidenciaram o que viria a ser

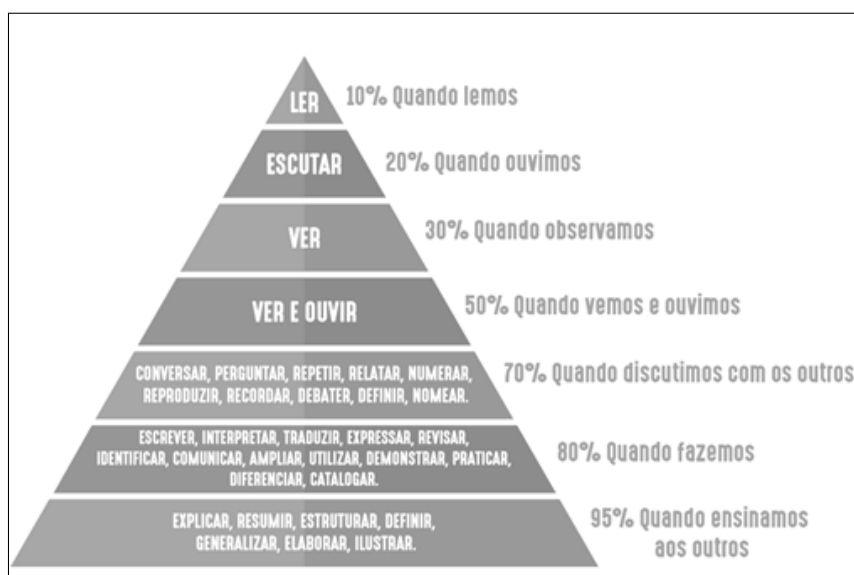
uma grande descoberta no meio educacional, em condições iguais de ensino (ignorando as variáveis externas à atmosfera educacional) todos os alunos aprendiam, mas se diferenciavam em relação ao nível de profundidade e subjetividade do conhecimento adquirido.

Dentro dessa realidade, visando a melhora na absorção do conhecimento, têm surgido diversas propostas de práticas pedagógicas alternativas como por exemplo, a aprendizagem e/ou metodologia ativa. A aprendizagem ativa baseia-se na transmissão de informação, onde o aluno assume uma postura mais participativa na qual ele resolve problemas, desenvolve projetos e cria oportunidades para a construção de conhecimento, possibilitando ensinar e aprender de modo participativo.

O estudante precisa assumir um papel cada vez mais ativo, descondicionado da atitude de mero receptor de conteúdos, buscando efetivamente conhecimentos relevantes aos problemas e aos objetivos da aprendizagem. Iniciativa criadora, curiosidade científica, espírito crítico reflexivo, capacidade para auto-avaliação, cooperação para o trabalho em equipe, senso de responsabilidade, ética e sensibilidade na assistência são características fundamentais a ser desenvolvidas em seu perfil. (MITRE *et al.*, 2008).

O psiquiatra americano William Glasser desenvolveu um estudo onde os níveis de assimilação do conhecimento são apresentados. A Figura 1 mostra a pirâmide elaborada pelo autor que retrata a forma como aprendemos. Na opinião de Gomes *et al.* (2018) a pesquisa destaca, ainda, a necessidade de propor atividades que não trabalhem apenas a memorização de conteúdos como forma de aprendizagem.

Figura 1 – Pirâmide de Aprendizagem de William Glasse



Fonte: Adaptado de Nunes e Bessa (2018)

Em meados de 1960, David Ausubel formulou a Teoria da Aprendizagem Significativa, onde destacou a aprendizagem de significados (conceitos) como a mais importante

forma de aprendizagem. Ele enfatizou que boa parte da aprendizagem ocorre de forma receptiva e, com isso, a humanidade tem-se valido para transmitir as informações ao longo das gerações.

Nesse contexto, vários métodos surgiram para apoiar o uso da aprendizagem significativa, um deles foi descrito por Barbosa e Moura (2013), eles afirmaram que mapas conceituais caracterizam-se como artifício ativo que faz com que os alunos relacionem o conteúdo e assimilem o conhecimento, conectando novos saberes com os que já possuem, o que possibilita uma aprendizagem mais significativa.

Assim sendo, este trabalho descreve o desenvolvimento de uma arquitetura computacional que emprega os conceitos de aprendizagem significativa e o uso de mapas conceituais para identificar o nível de cognição do aprendiz, utilizando Processamento de Linguagem Natural (PLN).

1.2 PROBLEMÁTICA

As afirmações contidas nessa pesquisa, assim como o motivo da sua investigação, são compreendidas à luz da Educação. São ações que buscam opções favoráveis à inserção de diferentes tecnologias e não apenas exercícios computacionais, tudo isso com o objetivo de criar possibilidades para a descoberta de diferentes formas de ensinar e aprender.

Some-se a isso o papel da aula eficaz e do gerenciamento de tempo, aqui relatado por Puentes e Aquino (2008) que pressupõe a obtenção de bons resultados e alta produtividade no rendimento do trabalho do aluno e do professor, com as inversões de tempo e esforços indispensáveis.

Acrescenta-se, também, que o desenvolvimento, avaliação e evolução dos mapas conceituais conduz para análise do grau de conhecimento do aluno sobre um determinado tema, no qual esse princípio já vem sendo estudado a alguns anos. Além disso, a possibilidade de medir esse conhecimento de forma automática, com a utilização de ferramentas computacionais que possibilitem a otimização do tempo efetivo de sala aula, nos instigam a pesquisar sobre o assunto.

Por sua vez, o âmbito dessa pesquisa limitou-se às abordagens de construção de mapas conceituais na visão de Novak e Cañas (2010), no processamento textual das proposições através de regras gramaticais de sintática e semântica usando PLN e na geração de scores para a estrutura geral dos mapas conceituais de acordo com o que teoriza Cañas *et al.* (2006).

Pelas razões aqui reveladas, surgiu a seguinte hipótese: Como inferir, de forma automática, o nível cognitivo do aluno analisando os dados extraído de mapas conceituais, fazendo uso das proposições textuais obtidas dessa extração por intermédio de

Processamento de Linguagem Natural (PLN)?

1.3 OBJETIVOS

1.3.1 Objetivo Geral

Desenvolver uma Arquitetura Computacional que utilize recurso de Processamento de Linguagem Natural para classificar Mapas Conceituais elaborados por aprendizes, permitindo a sua diferenciação cognitiva.

1.3.2 Objetivos Específicos

Os objetivos específicos descrevem, com mais detalhes, os temas necessários ao objeto da pesquisa, visto que, esses argumentos norteiam os conteúdos que serão apresentados nesse estudo. Para realização do trabalho desenvolveu-se os seguintes objetivos:

- I. Discutir a Teoria da Aprendizagem Significativa;
- II. Levantar os requisitos de criação de Mapas Conceituais;
- III. Realizar a extração das proposições dos Mapas Conceituais;
- IV. Identificar as características gramaticais e contextuais das proposições com auxílio de ferramentas de PLN;
- V. Validar as Proposições;
- VI. Implementar Sistema de Scores;
- VII. Realizar Análise Topológica;
- VIII. Validar o Sistema com dados reais.

1.4 METODOLOGIA

Segundo Moresi *et al.* (2003), é nessa fase que se define como e onde a pesquisa será realizada, e como será categorizada. Nesse estudo utiliza-se a seguinte classificação:

- Para os objetivos desse trabalho, aplica-se a pesquisa bibliográfica para embasar a teoria da aprendizagem significativa auxiliando a desenvolver, explicar e modificar conceitos e ideias;

- No que respeita os critérios técnicos, adota-se a pesquisa exploratória visando auxiliar o trabalho com Mapas Conceituais e o Processamento de Linguagem Natural, se propondo a realizar uma análise da literatura existente.

Em outras palavras, nosso estudo foi baseado em pesquisa exploratória com o objetivo de buscar metodologias, conceitos e técnicas voltadas a aplicações que estudem, de alguma forma, Mapas Conceituais com eixo em educação utilizando Processamento de Linguagem Natural.

Do mesmo modo, procurou-se identificar a real utilidade de desenvolver uma ferramenta voltada para o professor, que consiga classificar o nível de cognição do aprendiz utilizando como base a Teoria da Aprendizagem Significativa.

1.5 MOTIVAÇÃO

Segundo Ausubel, Novak e Hanesian (1980), O fator isolado mais importante que influencia a aprendizagem é aquilo que o aprendiz já conhece. Descubra o que o aluno sabe e baseando-se nisso seus ensinamentos.

Apoiado nessa afirmação, Novak e Cañas (2010) buscou representar esse aprendizado guardado e organizado na estrutura cognitiva do aluno de uma forma mais didática. Diante dessa provocação, o autor desenvolveu a metodologia de Mapas Conceituais, onde o professor poderá identificar como o aprendiz conhece os conceitos, estrutura-os, relaciona com outros conceitos, discrimina e integra baseado em um conteúdo ou disciplina específicos, conforme explica Moreira (2013).

Em vista disso, faz-se necessário desenvolver um método que consiga reunir a habilidade de compreender o que o discente já conhece sobre o conteúdo e a tradução desse conhecimento impresso nos mapas conceituais. Ao mesmo tempo, as Tecnologias Digitais de Informação e Comunicação (TDIC) vem auxiliar esse trabalho, possibilitando a criação de novas soluções que unam professores, alunos, pessoas, objetos e informações que estejam envolvidos nesse processo.

1.6 ORGANIZAÇÃO DO DOCUMENTO

Este trabalho está organizado da seguinte forma: o Capítulo 2 apresenta o Referencial Teórico, tratando de temas como: Arquitetura de Sistemas, Teoria da Aprendizagem Significativa, Diferenciação Progressiva, Reconciliação Integrativa, Mapas Conceituais, Conceitos Obrigatórios, Avaliação de Mapas Conceituais, Taxonomia Topológica, *Scoring Criteria for Concept Maps*, Processamento de Linguagem Natural, Bibliotecas NLTK e spaCy. O Capítulo 3 vem abordando os trabalhos relacionados com os temas propostos

anteriormente, com objetivo de justificar a pesquisa. Em seguida, o Capítulo 4 traz a construção da Arquitetura Computacional, evidenciando todos os seus módulos e, em seguida, a sua implementação com as respectivas validações em um cenário real. Por fim, o Capítulo 5 encerra a pesquisa com as Conclusões, limitações e Trabalhos Futuros.

2 Aspectos Conceituais

Este Capítulo estabelece as diretrizes conceituais da pesquisa, e servirá de alicerce para a construção do conhecimento necessário. Segundo Moresi *et al.* (2003), a revisão de literatura é o resultado do levantamento e pesquisa das publicações referidas ao tema em questão.

O nosso estudo utiliza a revisão teórica, referenciando e comparando os principais conceitos norteadores da pesquisa. Em síntese o trabalho aborda os princípios da Teoria da Aprendizagem Significativa, procurando explorar a relação entre o conhecimento prévio e o novo conhecimento, buscando ainda explicar de que forma esses conceitos se relacionam e se reconectam. Além disso, associa a ideia dos conceitos basilares da utilização de Mapas Conceituais, fazendo a relação dos seus métodos de criação e estruturação com a Aprendizagem Significativa. E, por fim, apresentaremos as técnicas de Processamento de Linguagem Natural (PLN) voltadas ao reconhecimento de padrões linguísticos que ajudem a indetificar as relações gramaticais nas proposições geradas pelos Mapas Conceituais, possibilitando assim classificar esses e auxiliar o professor no processo de ensino aprendizagem.

2.1 TEORIAS DA APRENDIZAGEM SIGNIFICATIVA

O conhecimento proporciona liberdade, integra o indivíduo com o seu meio no que tange o respeito as regras, princípios, sentimentos e atitudes. Com a independência da sua estrutura de conhecimento, o indivíduo se torna capacitado a apreender outros cenários semelhantes de conhecimento, tornando-se assim, dono da informação. Para Vygotsky (1988), todas as funções mentais superiores têm origem com as relações entre seres humanos.

Nessa perspectiva as relações sociais estão vinculadas, de modo direto, a aquisição de significados. A definição das palavras e dos gestos são adquiridos socialmente, e nesse sentido produzem relações contextuais, conforme teoriza Moreira (2009). O processo de ensino e de aprendizagem atingiu singular importância quando David Ausubel, em meados de 1963, apresentou a comunidade científica, a concepção de um novo modelo de aprendizagem, a significativa, focada no conhecimento adquirido. Aquela época, tinha-se a ideia de aprendizagem e de não aprendizagem, conceitos que perduraram por longos anos.

Segundo Braathen (2012), o processo de ensino-aprendizagem envolve três questões basilares, sendo o ensino, a aprendizagem e a avaliação. Há de se destacar a importância da compreensão desses aspectos pelo professor para que o processo seja válido. A Figura 2 exibe os elementos do processo ensino-aprendizagem, o qual para se ensinar com excelência

há a necessidade de aprender de igual forma, por esse motivo é que a aprendizagem está em primeiro lugar, mas não dissociada do ensino.

Figura 2 – Elementos do processo de ensino-aprendizagem



Fonte: Adaptado de Braathen (2012)

A espinha dorsal da teoria da aprendizagem significativa de Ausubel, Novak e Hanesian (1980) refere-se a importância dada à aquilo que o indivíduo já conhece, do conhecimento já assimilado, ou seja, o conhecimento já existente em sua estrutura cognitiva com compressão, equilíbrio e diferenciação, esse conhecimento é chamado pela literatura de subsunçores, sendo necessário, obviamente, investigar previamente tal entedimento para adotar estratégias de ensino adequadas.

Um processo através do qual uma nova informação se relaciona, de maneira substantiva (não-litera) e não-arbitrária, a um aspecto relevante da estrutura cognitiva do indivíduo. Neste processo a nova informação interage com uma estrutura de conhecimento específica, a qual Ausubel chama de "conceito subsunçor" ou, simplesmente "subsunçor", existente na estrutura cognitiva de quem aprende (MOREIRA, 2009).

Permanecendo nesse contexto, a literatura elenca três condições elementares para se aprender de forma significativa: a apresentação de um novo conhecimento formulado de maneira lógica; a presença de conceitos prévios na estrutura cognitiva, que possibilite a sua relação com o novo conhecimento; a ação clara de apreender e unir o conhecimento prévio com o que pretende absorver. Esses conhecimentos antecipados são também chamados de conceitos subsunçores ou âncora.

Na visão humanista de Novak, quando a aprendizagem é significativa o aluno evolui, tem prazer no que aprende e é estimulado a buscar novas aprendizagens. Entretanto, quando o oposto ocorre, a aprendizagem é sempre mecânica e, nesse cenário, o indivíduo desenvolver verdadeira repulsa à matéria e não se predispõe à aprendizagem significativa.

2.1.1 Aprendizagem Significativa x Aprendizagem Mecânica

Em contraste com as características do estudo teorizado por Ausubel *et al.* (1968), usa-se um modelo onde a aprendizagem é carente de correlação com o conhecimento

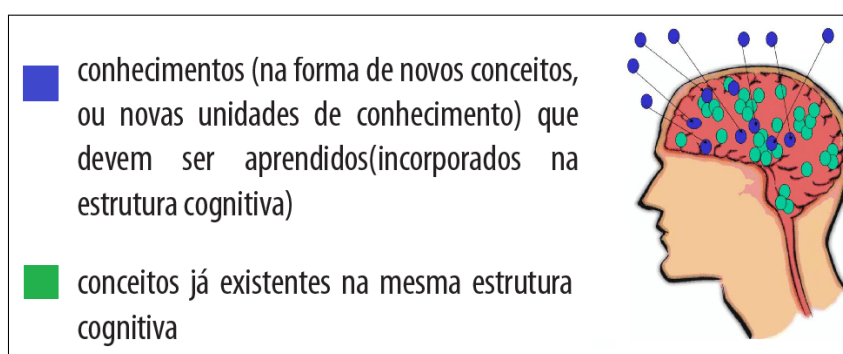
prévio do aprendiz e, inexistente uma atribuição de significados pessoais àquele conhecimento, esse cenário não é considerado como sendo aprendizagem significativa e sim mecânica, conhecida no Brasil como “decoreba”.

Novas informações são aprendidas praticamente sem interagirem com conceitos relevantes existentes na estrutura cognitiva, sem ligarem-se a conceitos subsuntores específicos. A nova informação é armazenada de maneira arbitrária e literal, não interagindo com aquela já existente na estrutura cognitiva e pouco ou nada contribuindo para sua elaboração e diferenciação (MOREIRA, 2009).

A Aprendizagem Mecânica acontece quando o discente introduz um novo conhecimento de forma arbitrária, isso quer dizer que o discente aprende sem entender o real motivo ou, da mesma forma, não compreender o significado do porquê. Essa ação de incorporar um novo conhecimento também pode ocorrer de maneira literal, o aluno aprende rigorosamente como foi falado ou escrito, sem margem para uma compreensão mais particular. A aprendizagem decorre da inexistência de conhecimento prévio relativo ao novo conhecimento aprendido. Pode-se exemplificar esse contexto com o seguinte cenário: O aluno aprende sobre sistema de memórias de um hardware sem saber a diferença entre RAM e ROM ou mesmo o sistema binário computacional.

Ainda pensando sobre isso, ilustramos na Figura 3 o comportamento cerebral no processo de aprendizagem segundo Braathen (2012), onde o conhecimento agrupado representa a aprendizagem significativa e quando esse conhecimento se encontra isolado, temos a mecânica.

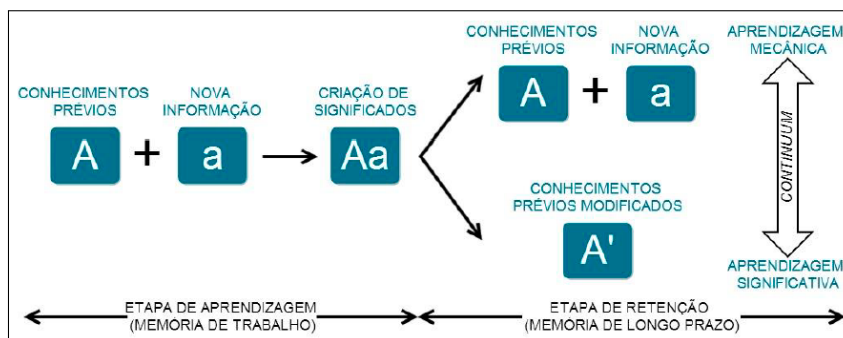
Figura 3 – Ilustração do processo de aprendizagem Mecânica e de Aprendizagem significativa



Fonte: Adaptado de Braathen (2012)

Continuando nesse tema, a Figura 4 apresenta de forma mais didática a trajetória das relações conceituais entre a aprendizagem significativa e mecânica. O conhecimento prévio e novo conhecimento inicialmente são armazenados na memória de trabalho (memória de curto prazo), após a geração de significado há uma fusão desses dois conhecimentos na memória de longo prazo, produzindo assim a aprendizagem significativa, o contrário ocorre com a aprendizagem mecânica, que continua dissociada na memória de longo prazo.

Figura 4 – Comparação entre aprendizagem mecânica e significativa de acordo com a Teoria de Ausubel



Fonte: Aguiar e Correia (2013)

A aprendizagem significativa propõe uma metodologia mais simples, funcional e eficaz de armazenar o conhecimento, já a aprendizagem mecânica requer um tempo maior e um esforço moderado para assimilar conceitos. Esse custo maior poderia ser simplificado caso o aprendiz conseguisse associar um subsunçor existente na estrutura cognitiva a esse novo conceito. Desta forma, Ausubel recomenda que a programação do conteúdo a ser ensinado siga dois princípios básicos: a diferenciação progressiva e a reconciliação integrativa.

2.1.2 Diferenciação Progressiva

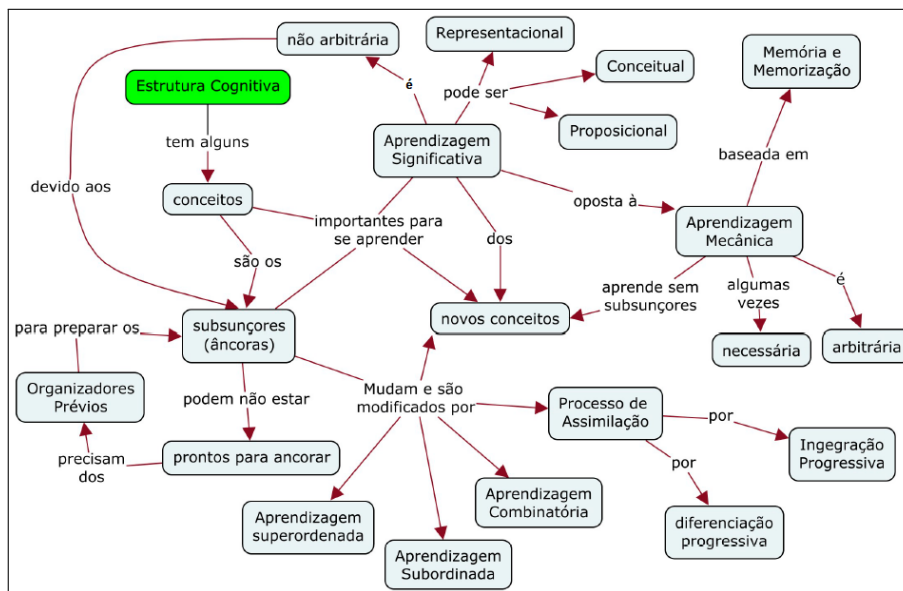
A literatura descreve o conceito da diferenciação progressiva como sendo o elemento necessário na preparação das aulas e programação do conteúdo. Essa premissa foi descrita por Moreira (2009) como “o princípio segundo o qual as idéias e conceitos mais gerais e inclusivos do conteúdo da matéria de ensino devem ser apresentados no início da instrução e, progressivamente, diferenciados em termos de detalhe e especificidade”, nesse contexto as ideias mais genéricas devem ser organizadas no topo da estrutura cognitiva do aluno e na medida em que esse aluno avança no conteúdo, ele vai recepcionando e incluindo progressivamente conceitos mais específicos, aprofundando assim esses relacionamentos.

2.1.3 Reconciliação Integrativa

Do mesmo modo, o referido autor destaca o princípio da reconciliação integrativa como segundo elemento necessário para a aprendizagem significativa, segundo ele tal princípio destaca a relação entre ideias, indicando semelhanças e diferenças importantes, reconciliando discrepâncias reais ou aparentes. Em outras palavras, trata-se da relação que o aluno estabelece dos novos conceitos com os já preexistentes na estrutura cognitiva, com isso é possível recombina-los já existente, reorganiza-los e construir novos significados.

Para fins didáticos, apresentamos na Figura 5 uma ilustração organizada conforme o entendimento da teoria da aprendizagem significativa de Ausubel, esse modelo utilizou-se da ideia de mapas conceituais para visualização da referida teoria.

Figura 5 – Uma versão reduzida da teoria de aprendizagem de Ausubel



Fonte: Filho e Ferreira (2018)

Mediante o exposto, uma vez adquirindo a aprendizagem significativa, o aluno converte a interpretação lógica do material pedagógico em uma definição psicológica e, nesse momento, o conteúdo é inserido de forma bem específica na sua estrutura cognitiva, sendo bem peculiar essa inserção em cada indivíduo, o que torna essa atitude um processo idiossincrático. Esse método foi representado por Joseph Novak na década de 1970, utilizando organizadores gráficos que conseguiram estruturar a organização do conhecimento em conceitos e proposições.

2.2 MAPAS CONCEITUAIS

A ideia original dos Mapas Conceituais surge da inquietação do pesquisador norte-americano Joseph Novak em investigar os fatores limitadores do aprendizado em crianças de seis a oito anos, o estudo tinha a pretensão de pesquisar a construção da estrutura cognitiva da criança no processo de aprendizagem.

À medida que transcrevíamos as fitas, nós pudemos observar que as proposições usadas pelos estudantes normalmente melhoravam em relevância, número e qualidade, mas ainda era difícil observar especificamente como suas estruturas cognitivas estavam mudando (NOVAK; CAÑAS, 2006a).

Levando em consideração as limitações dos métodos de pesquisa à época, surgiu a necessidade de buscar alternativas mais robustas que pudessem melhorar a análise

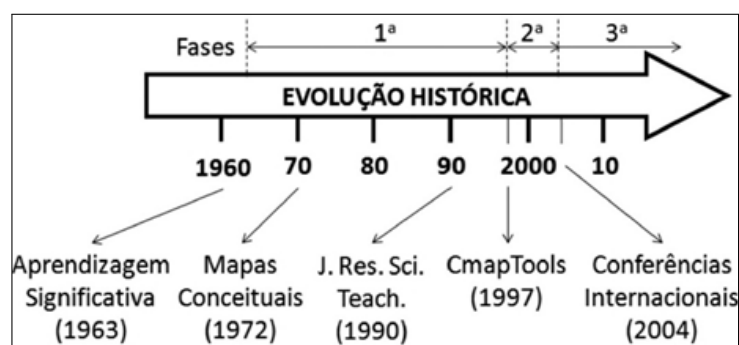
das informações. Diante desse contexto, metodologias diagramáticas apoiadas na teoria cognitivista de Ausubel apontavam como alternativa, e assim os estudos evoluíram e os resultados geraram o que conhecemos hoje como mapas conceituais.

Salienta-se ainda que a essência do estudo sobre mapas conceituais está correlacionado com a teoria da aprendizagem significativa, isso implica em um empenho maior do aprendiz em resgatar os conceitos adequados, já armazenados em sua estrutura intelectual, para conecta-los com as novas informações, em vista disso a estrutura cognitiva torna-se alicerces para o aprendizado. De certo que o estudo sobre mapas conceituais e a teoria da aprendizagem significativa não foi um evento casual, Novak intencionalmente estruturou sua pesquisa embasada na teoria de Ausubel.

Diante disso, tem-se a ideia de mapa conceitual como sendo uma representação idiossincrática do autor com foco em seus relacionamentos. Quando o professor desenvolve um mapa ele expressa o seu conhecimento completo sobre um tema, da mesma forma, o aprendiz aprimora o seu know-how, construindo uma hierarquia conceitual.

Conforme afirma Kowata (2010), o conceito de “mapas conceituais” é usado sem moderação pela literatura, onde suas diversas especificações e uso confundem sua correta utilização no decorrer da história. Da mesma forma essas variantes acabam levando a criação de novos sentidos ao que decerto temos como teoria original. A Figura 6 ilustra como se deu a evolução histórica do mapeamento conceitual através da linha do tempo que se inicia em meados da década de 60 com a Teoria da Aprendizagem Significativa e chegando até as conferências internacionais.

Figura 6 – Evolução do mapeamento conceitual num contexto histórico

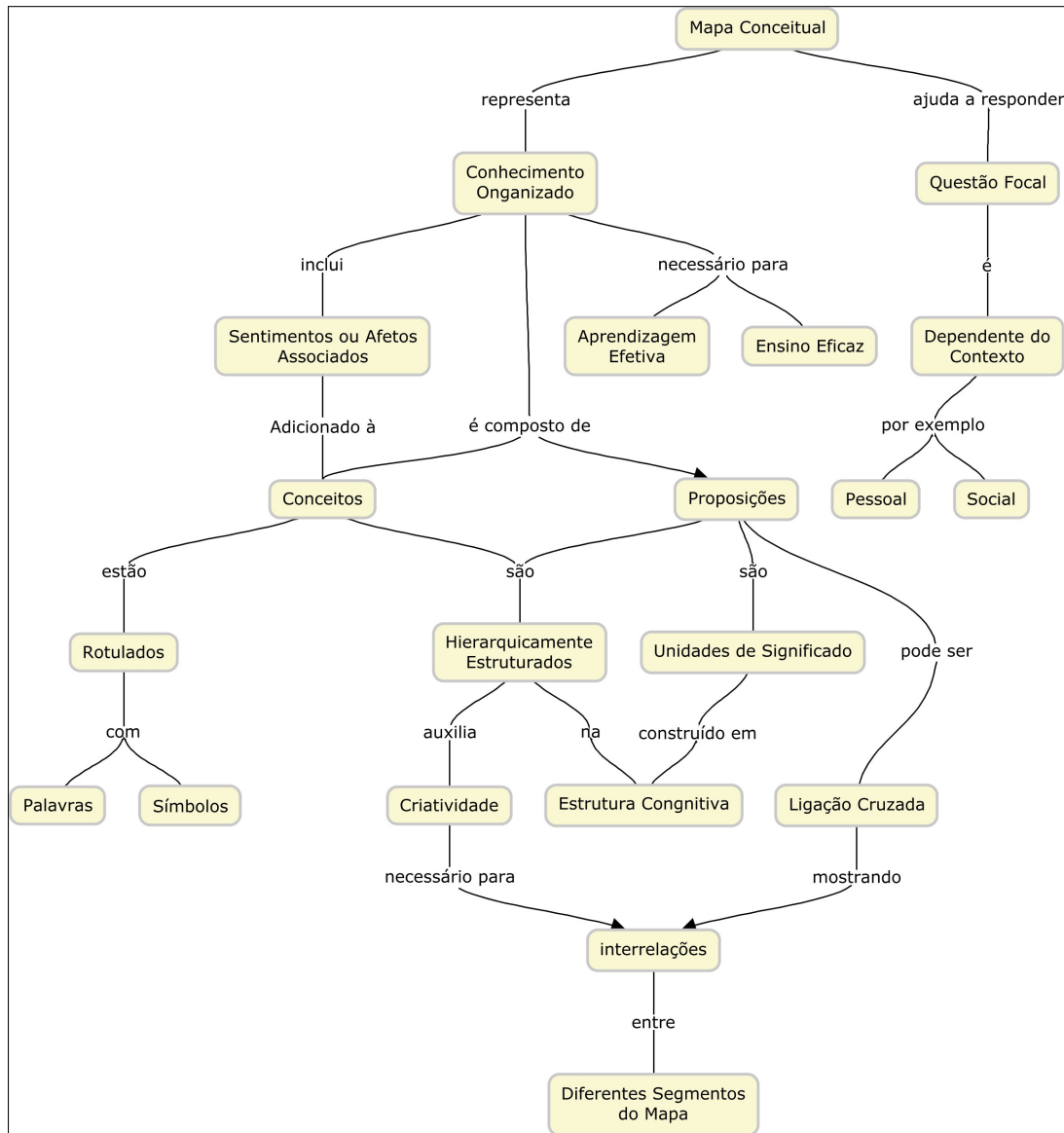


Fonte: Cicuto, Correia *et al.* (2012)

Para esse trabalho, usaremos a concepção de mapas conceituais vinculadas a forma de diagramas, obedecendo modelo de especificação válido derivado da linguagem de especificação. Essas especificações são baseadas em símbolos pictográficos, caracterizados por retângulos e segmentos de reta, os dois devidamente rotulados que, por conseguinte, formam proposições com sentido lógico representando assim a “pergunta focal”, conforme apresenta a Figura 7 que exhibe a estrutura de criação de Mapas Conceituais segundo Novak

e Cañas (2006b). Essa instância representa evento importante para concepção do mapa conceitual, traduzindo-se em um acontecimento ou fenômeno pelo qual se deseja entender seu real sentido, iniciando com isso a hierarquização da estrutura das ideias dentro de um contexto.

Figura 7 – Estrutura de Criação de um Mapa Conceitual



Fonte: Adaptado de Novak e Cañas (2006b)

De posse dessas variantes, convém lembrar que esses organizadores gráficos são instrumentos indicados para reproduzir o conhecimento, contribuindo para a permanência e/ou resgate das informações durante o processo de aprendizagem, conforme menciona Vekiri (2002). Esse resgate é explicado pelo processamento da informação na memória de trabalho, fenômeno referenciado no estudo de Sadoski e Paivio (2013), em sua Teoria da Dupla Codificação que esclarece que os impulsos verbais (palavras, conceitos, textos) e imagéticos (figuras, arranjos espaciais, cores) podem ser processados em paralelo na

memória de trabalho, sem provocar sobrecarga.

Tofilon e Fike (2000) afirmam ainda que a Teoria da Dupla Codificação (TDC), distingue três tipos de processamento: (i) representacional, a ativação direta de representações verbais ou não verbais; (ii) referencial, a ativação do sistema verbal pelo sistema não-verbal ou vice-versa; e (iii) o processamento associativo, ou ativação de representações dentro do mesmo sistema verbal ou não verbal. Para entender melhor, imagine um avião, forme a imagem mental desse avião em sua mente ou pense na palavra “avião”, dessa forma os sistemas verbais e de imagem estão associados, como se pode pensar na imagem mental do avião e em seguida representa-lo em palavras, ou ler ou ouvir palavras e depois formar uma imagem mental.

A fim de compreender como os elementos da TDC podem estar relacionados com a teoria alusiva a Mapas Conceituais, elencamos os pontos que devem ser abordados na construção de bons diagramas, conforme instrui Ruiz-Moreno *et al.* (2007) em seu estudo:

- I. **Elaboração da Pergunta Focal:** A escolha acertada da questão focal revela bastante sobre a qualidade do mapa conceitual, ou seja, “uma pergunta que especifica claramente o problema ou questão que o mapa conceitual deve ajudar a resolver” Novak e Cañas (2006a);
- II. **Escolha dos Conceitos:** Qualidade e quantidade de conceitos exibidos e níveis de hierarquia conceitual, identificando os conceitos mais gerais e deles partindo para conceitos mais específicos, sempre estruturando-os de forma hierárquica. Reproduzindo as ideias e análises de Novak e Cañas, a base média de conceitos para a construção de um mapa bem estruturado varia em torno 15 a 25 conceitos e, para obter essa média, deve-se levar em consideração o nível de conhecimento do aluno e suas habilidades em desenvolver mapas conceituais. Os mesmos autores ainda relatam que não se deve utilizar trechos de textos ao invés de conceitos;

Nós incluímos isso porque, em nosso ambiente, a presença de pedaços de texto em um mapa conceitual é geralmente indicativa de conhecimento memorial, e, portanto, pobre, rígido e isolado. Pensamos que, pelo contrário, a capacidade de quebrar textos em conceitos é o ponto de partida obrigatório para o estabelecimento de múltiplos relacionamentos, inovadora e flexível entre ideias e, portanto, para a construção de estruturas cognitivas cada vez mais complexas e sofisticado. Atribuímos tanta importância a isso, que um mapa dominado por pedaços de texto recebe em nossa taxonomia a menor avaliação (0), independente de sua complexidade ou da presença de outro elemento estrutural (NOVAK; CAÑAS, 2006a) (tradução nossa e grifo nosso).

- III. **Palavras de Ligação entre os Conceitos:** Reproduzindo o estudo de Aguiar e Correia (2013), há que se considerar, portanto, que usar um termo de ligação que indica falta de clareza insinua uma compreensão incompleta dos alunos sobre o conteúdo

conceitual. O mesmo estudo complementa ainda afirmando que “*variações no tempo verbal do termo de ligação são suficientes para mudar drasticamente o seu sentido*”. Seguindo essa mesma perspectiva, as frases de ligação podem ser compostas por verbos na opinião de Novak e Gowin (1984);

- IV. Proposições: A junção dos conceitos e das palavras de ligação devem formar um sentido lógico, levando a uma unidade de significado.

Consoante com essa abordagem, dizemos que o estudo sobre Mapas Conceituais pode apontar informações de potencial significativo sobre um domínio específico quando analisada de uma ótica mais detalhista, o contrário não ocorreria se essa mesma análise fosse apenas mera leitura de suas proposições. Ancorado nessa ideia de análise mais estruturada em um cenário de proposições e, levando em consideração a sobrecarga de trabalho inputada ao professor no cenário atual da docência, é conveniente pensar em uma forma de processar toda essa informação com o objetivo de extrair representações e significados mais tangíveis. Segundo Gomes (2019), os estudos na área de mineração de dados apontam para um aumento da pesquisa acerca do tema Processamento de Linguagem Natural (PLN) como forma de resolver problemas semelhantes.

2.3 PROCESSAMENTO DE LINGUAGEM NATURAL

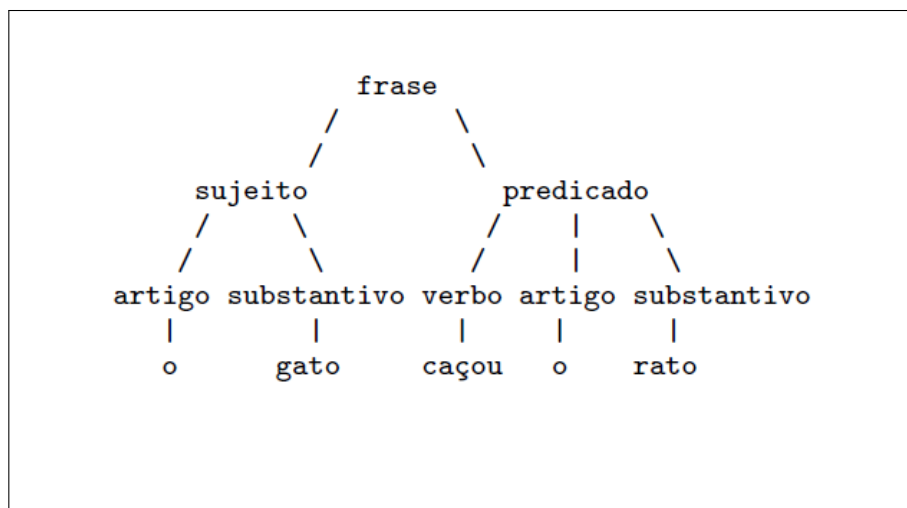
Processamento de Linguagem Natural (PLN) é uma subárea da inteligência Artificial (IA) responsável por criar modelos computacionais que desempenham tarefas apoiadas em linguagem natural. Segundo Young *et al.* (2018), atualmente as pesquisas em PLN estão concentradas em abordagens que trabalhem com aprendizagem profunda (*deep learning*). Essas abordagens facilitam o processamento ágil de grandes volumes de dados textuais, pois tem como propósito minerar informações significativas buscando padrões para classificação de textos. Por outro lado, Rizkallah (2017) afirma que uma vez que esse volume enorme de dados representa uma oportunidade para as corporações, a carga computacional utilizada para extrair informações importantes torna-se um desafio que muitas vezes é negligenciado por essas mesmas organizações.

Do mesmo modo a comunidade científica presencia um crescimento significativo no volume de informações geradas e armazenadas de forma não sistematizada, ou seja, dados que não possuem uma indexação que facilite a sua recuperação.

Algoritmos de PLN representam modelos matemáticos voltados ao processamento textual, utilizando-se de bases de dados contextualizada (*corpora*) para aprender computacionalmente. Segundo Gomes (2019), esse entendimento reflete a tentativa de processar e associar características fundamentais da linguagem com: morfologia, sintaxe e semântica.

A morfologia estuda a arrumação, a composição e a classificação das palavras, reconhecendo as palavras pela composição das suas unidades primitivas. Sintaxe considera as regras que constroem as frases nas línguas naturais, ou seja, a parte da gramática que estuda a arrumação das palavras na frase e essa no discurso. A semântica discute o significado e o atribui as palavras em uma estrutura sintática conforme visualizamos na Figura 8.

Figura 8 – Árvore Sintática



Fonte: Pereira (2011)

Com a utilização dessas grandes bases de dados textuais, pode-se utilizar metodologias e cálculos de diversos tipos de grandezas para obter informação úteis como sequência, ocorrência e características de palavras. Essas bases quando sistematizadas são denominadas de corpus textuais.

Para Carvalho (2012) o corpus textual é um repositório *sui generis*, onde são coletados textos conforme critérios específicos, empregado em treinamento de modelos estatísticos em linguagem natural. Algumas técnicas são utilizadas com um padrão especial de corpus textual chamado de corpus anotado, esse utiliza estruturas básicas de marcação como delimitação dos parágrafos e das orações, estrutura sintática e categorização das palavras. A Tabela 1 mostra um fragmento de texto extraído da obra de Bandeira (1964) e em seguida o mesmo texto com as marcações morfológicas.

Além disso, algumas técnicas podem ser adicionadas para que a extração da informação tenha uma acurácia maior, à citar a normalização de texto, a segmentação, tokenização e análise lexical.

- Normalização de Texto: Essa técnica consiste em remover os marcadores de formatação (tags, estilos de fontes e metadados) presentes no conteúdo do documento, essa extração facilita a manipulação de arquivos em diferentes formatos;

Tabela 1 – Comparação entre texto normal e texto marcado

Texto sem Marcação	Eu disponho para vós o Reino, como meu Pai o dispôs para mim;
Texto Marcado	Eu/PRO disponho/VB-P para/P vós/PRO o/D Reino/NPR ,/, como/CONJS meu/PRO Pai/NPR o/CL dispôs/VB-D para/P mim/PRO ;/.

Fonte: Adaptado de Galves e Faria (2010)

- Segmentação: Segundo Carvalho (2012), trata-se da divisão do conteúdo do texto em sentenças. Um grande problema encontrado nessa técnica é identificar corretamente os caracteres finalizadores de sentenças, tais como: “?”, “:”, “;”, “/” e “.” onde, igualmente, esses sinais também podem definir abreviações de nomes, separação de dígitos em datas, horas, números de telefones e números ordinais ;
- *Tokenização*: São os elementos que constituem a oração, palavras, números e sinais de pontuação e espaços, esses elementos definem a menor unidade que pode ser identificada dentro do cenário da linguagem natural. Por outro lado, Pereira (2014) afirmou que apenas a tokenização pode não alcançar uma boa acurácia, dessa forma aconselha-se a utilização concomitante dessa técnica com o reconhecimento de nomes de entidades;
- Análise Lexical: Conforme explica Kowata (2010), essa etapa atribui a cada token uma tag morfológica correlata. A análise lexical também é conhecida como classe morfológica, etiqueta lexical, classe de palavras e POS e distingue substantivo, verbo, pronome, preposição, advérbio, conjunção ou artigo. Em seguida é retornado um arquivo contendo tokens anotados por suas respectivas etiquetas morfológicas.

Uma das dificuldades desta tarefa é a existência de muitas palavras com diferentes classificações possíveis. Tais palavras, se estiverem fora de contexto, ocasionam a ambiguidade sobre sua interpretação para a correta classificação gramatical (CARVALHO, 2012).

2.4 PROGRAMAÇÃO PARA PLN

Em adição aos assuntos já mencionados sobre PLN, cita-se de forma adicional os *frameworks* utilizados nessa pesquisa, que serviram de embasamento para os algoritmos desenvolvidos.

2.4.1 Biblioteca NLTK

A *Natural Language Toolkit* (NLTK) é uma biblioteca de código aberto que trabalha com linguagem natural oferecendo diversos recursos linguísticos. Muito utilizada no meio

acadêmico o NLTK é desenvolvido em Python e trabalha diretamente com *tokenização* de palavras e frases, *stemmers*, classificadores, *part-of-speech taggers* e *chunkers*. Sua documentação é robusta e relativamente fácil de aprender e utilizar.

Dentre as diversas funcionalidades do NLTK, sobressai como ponto forte em nossa pesquisa a segmentação de sentenças, ação caracterizada pelo reconhecimento de marcadores que delimitam as orações, sendo esses os sinais de pontuação. Segundo Kowata, Cury e Boeres (2012) um obstáculo a essa ação é a correta vinculação dos sinais de pontuação ao fim de sentença, haja vista que esses caracteres finalizadores também podem especificar abreviações, separações de dígitos de datas, horas e outras representações numéricas. Para resolver esse problema associa-se a esse método a *tokenização*, remoção de *stopwords* e a redução ao *stemming*.

Ainda sobre o *framework* destaca-se o processo de *Tokenização*, é nesse momento que as orações são divididas em *tokens* (palavras), usando sinais gráficos como pontos limítrofes, essa ação prepara o *token* para sua identificação gramatical, etapa executada pela biblioteca spaCy¹, que coloca uma etiqueta morfológica no *token*, identificando-o como substantivo, verbo, pronome, preposição, advérbio, conjunção, artigos e etc. Em estudos envolvendo análise de Mapas Conceituais, as duas técnicas são utilizadas para identificar os verbos nas proposições, isso possibilita o reconhecimento correto da estrutura da oração.

2.4.2 Biblioteca spaCy

Assim como o NLTK, o spaCy também é um *framework* avançado voltado para PLN, suporta mais de 33 línguas, dentre elas o português. Recomendado para trabalhos com muito texto natural, projetado para utilização em ambiente de produção que utilizam extração de informações, compreensão de linguagem natural e aprendizado profundo. Suas funcionalidades mais relevantes são: *Tokenização*, Marcação de parte da fala, Análise de Dependência, Lematização, Detecção de Limite de Frases, Semelhança e Treinamento.

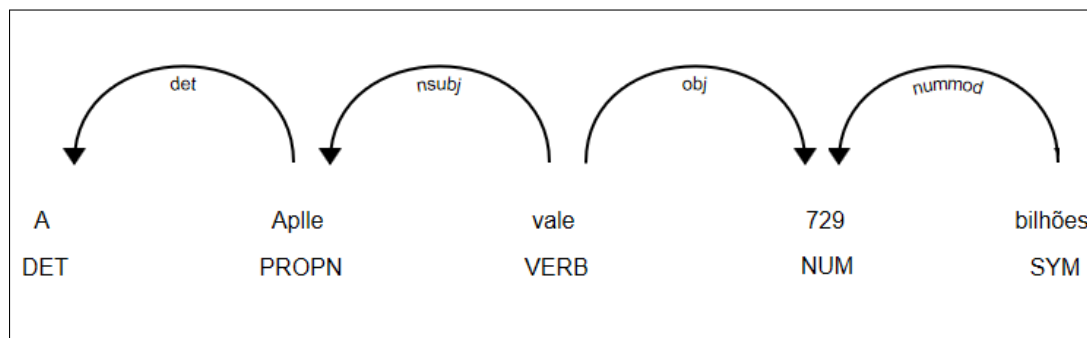
A classificação gramatical é o momento em que os *tokens* são rotulados, esse processamento é mais complexo, haja vista o contexto do cenário. A simples mudança na disposição das palavras em uma frase pode mudar o sentido da oração, e com isso provocar classificações diferentes dos *tokens*. Em um ambiente de análise proposicional, onde as frases são retiradas de Mapas Conceituais, a forma como a oração é construída e as palavras utilizadas nessa composição é condição necessária para um correto entendimento do mapa.

Para isso, o spaCy utiliza conhecimentos linguísticos mais apurados baseados em modelos estatísticos, onde o texto é analisado e o *framework* retorna um objeto com uma

¹ <https://spacy.io/>

variedade de anotações. Visualiza-se na Figura 9 como o spaCy rotula os *tokens* após processamento, é possível verificar a classificação de cada *token* e a relação de dependência entre eles, na Figura 10 o *framework* identifica as entidades contidas na oração e as rotula.

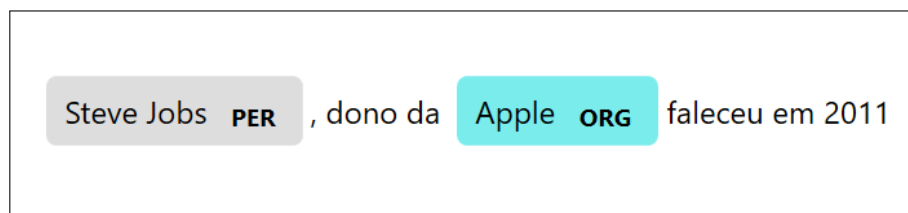
Figura 9 – Análise de Dependência Sintática



Fonte: Autoria Própria

Para um entendimento mais claro da análise de dependência, identificamos os rótulos que aparecem na Figura 9: **DET**: É o termo que acompanha o substantivo. **NSUBJ**: O sujeito do Verbo. **OBJ**: Objeto que complementa o sentido do verbo. **NUMMOD**: Qualificador numérico.

Figura 10 – Reconhecimento de Entidade Nomeada (REN)



Fonte: Autoria Própria

Na Figura 10 a identificação das entidades é realizada através de modelos treinados no corpus da Wikipédia, a classificação utiliza 7200 artigos da enciclopédia e os rotula manualmente em nove idiomas, atingindo uma precisão de 95%, conforme afirma Nothman *et al.* (2013). Na ilustração o modelo reconhece a entidade **PER**, que significa pessoa ou família nomeada e **ORG** que expressa uma entidade corporativa, governamental ou outra entidade organizacional.

Resumidamente, em um ambiente de Aprendizagem Significativa que utiliza Mapas Conceituais, a etapa de identificação da estrutura gramatical propicia a validação das proposições, encontrando seus elementos de composição e identificando se esses elementos estão classificados de forma correta.

3 Trabalhos Relacionados

Esta seção apresenta, de forma sintetizada, os trabalhos mais relevantes que serviram de base para a formulação da nossa pesquisa. É importante destacar que essa escolha foi baseada em três critérios objetivos. O primeiro pressupõe o próprio objeto do trabalho, encontrar fundamentação teórica que auxilie a classificação de Mapas Conceituais, utilizando como base os conceitos da Aprendizagem Significativa com o auxílio de Processamento de Linguagem Natural. O segundo refere-se ao tema pesquisado, ou seja, estruturamos todo o conteúdo do trabalho em grupos e buscamos, para cada item, os estudos mais relevantes e, de forma paralela, apresentamos os trabalhos segundo sua cronologia. Vale lembrar que o terceiro critério só foi usado quando utilizamos mais de uma referência por grupo de assuntos. Após essa esquematização explicamos cada pesquisa e associamos seus resultados e contribuições para o conjunto desse estudo.

3.1 APRENDIZAGEM SIGNIFICATIVA

Estruturamos, para o assunto “Aprendizagem Significativa” os seguintes itens: Estruturação do Conhecimento e a Importância dos Subsúncos segundo Moreira (2006) e Lemos (2013), depois falamos sobre a Criação de Material de Apoio, e com isso mencionamos Braathen (2012) e, por fim, a relação da Teoria da Aprendizagem Significativa com Mapas Conceituais segundo os estudos de Jesus e Silva (2004), Martins, Linhares e Reis (2011) e Moraes, Santana e Viana-Barbosa (2011).

Moreira (2006) com um olhar mais teórico, historiou em sua pesquisa pontos de vista distintos sobre a Aprendizagem Significativa, transitando em diferentes cenários de forma cronológica, abordando a visão clássica de Ausubel (1963), Ausubel *et al.* (1968), perpassando pela teoria humanista de Novak (1981), indo até Gowin (1981), Gowin, Novak e Valadares (1996) com olhar interacionista, mais a frente referencia Johnson-Laird (1983) em sua visão cognitiva contemporânea, relata ainda a complexidade e progressividade de Vergnaud (1990), autopoietica de Maturana *et al.* (2001), computacional de Dorneles, Araujo e Veit (2006) até chegar a sua própria concepção sobre o tema. Em sua análise o referido pesquisador afirma que o conhecimento significativo deve ser adquirido de forma crítica com interação social.

O primeiro desses princípios implica a interação social e o questionamento como elementos centrais na facilitação da aprendizagem significativa crítica: é mais importante aprender a perguntar do que aprender "*respostas certas*"(MOREIRA, 2006).

O autor ainda expõe a forma tradicional de se ensinar, onde o professor escreve e o aluno copia, decora e reproduz, afirmando que essa abordagem deve ganhar outra roupagem e traçar estratégias mais ativas envolvendo de forma mais participativa os alunos.

Lemos (2013) procurou destacar em seu trabalho a associação da TAS com a rotina da prática docente. O pesquisador declara que a atividade de ensinar é complexa e dinâmica atuando em um ambiente social peculiar e pessoal do indivíduo, retratando suas relações afetivas e cognitivas. O trabalho argumenta, ainda, que o professor deve ter independência e liberdade para tomar suas decisões, que são específicas do fazer docente, fazendo alusão ao ensino ideal. Esse ensino ideal retrata o comprometimento do professor com a aprendizagem significativa do aluno, referenciando sua realidade e possibilitando entender e negociar significados, com isso o autor finaliza sua pesquisa afirmando que a estratégia de ensino deve ser aquela que *“valoriza a natureza do conhecimento que se deseja aprender, a natureza do conhecimento prévio do aluno bem como o seu perfil sócio afetivo, o contexto no qual ocorrerá o evento educativo, o tempo disponível para a sua realização”*.

Em nossa pesquisa, referencia-se esses trabalhos para entender como funciona a etapa de aquisição do conhecimento, momento em que se descobre os subsunçores e há a ancoragem dos novos conceitos, passo importante para se obter a Aprendizagem Significativa.

Cotinuando no assunto, Braathen (2012) abordou os três eixos do processo de ensino-aprendizagem (Aprendizagem – Ensino – Avaliação) apoiado na visão construtivista de (AUSUBEL; NOVAK; HANESIAN, 1980 apud VYGOTSKY, 1997), onde realiza um debate sobre a forma como o professor entende o aprendizado de alunos diferentes. Traz também resultados baseados em tutorias, necessário a estudantes que tenham conhecimento prévio deficiente. Ratificando que esses ambientes com interatividade e dinamismo podem diminuir a necessidade de assistência extraclasse.

Jesus e Silva (2004) realizaram um levantamento sobre o uso de organizadores prévios no ensino contextualizado de funções lineares. Esse trabalho aborda a utilização da Teoria da Aprendizagem Significativa – TAS no estudo da matemática, onde o autor detalha os tipos de aprendizagem utilizadas na psicologia educacional. A pesquisa sugere, ainda, a construção de organizadores prévios ou antecipatórios para o ensino de funções, conforme cita o referido autor *“Não deve ser um simples resumo do conteúdo que virá, mas um material em nível maior de generalização ou abrangência, devendo preparar o aluno para que o novo conceito seja ancorado em conceitos já existentes na sua estrutura cognitiva”*.

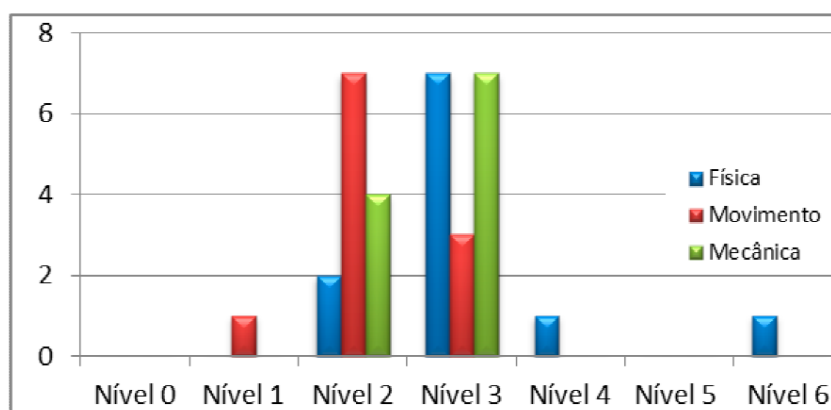
Martins, Linhares e Reis (2011) realizaram estudo com 42 alunos (nível fundamental e superior). Na pesquisa buscou-se identificar a reconciliação integrativa e diferenciação

progressiva dos significados e conceitos. Os resultados qualitativos foram obtidos através da comparação com um mapa de referência, onde as conclusões mostraram que 65% dos alunos adquiriram a compressão sobre o tema, resultado corroborado nas semelhanças das relações criadas em seus mapas com o mapa de referência. Esse estudo possibilitou perceber que os mapas conceituais são ferramentas capazes de identificar a aprendizagem significativa.

Com base nessas pesquisas consegue-se associar os conceitos de diferenciação progressiva e reconciliação integrativa no uso de Mapas Conceituais, isso possibilita não apenas classificar e pontuar Mapas Conceituais, como também associa-los a Aprendizagem Significativa. É importante ressaltar que a diferenciação progressiva aparece nos MP's no momento em que os conceitos mais genéricos, dispostos no topo do mapa, derivam para os mais específicos aprofundado o nível do Mapa Conceitual, e a reconciliação integrativa surge no instante em que galhos diferentes do mapa se unem em um nó específico.

Moraes, Santana e Viana-Barbosa (2011) apresentam um estudo, quantitativo e qualitativo, que relacionam a Aprendizagem Significativa com o uso de Mapas Conceituais. Da análise quantitativa gerou-se um *score* do mapa criado pelo aluno, tendo como base o mapa desenvolvido pelo professor, uma espécie de mapa de referência. Na análise qualitativa tomou-se como base a Taxonomia Topológica de Cañas *et al.* (2006). Como resultado da pesquisa, observou-se que a predominância de mapas nível 2 e nível 3, conforme aparece referenciado na Figura 11.

Figura 11 – Quantidade de mapas classificados por tema e por nível



Fonte: Adaptado de Moraes, Santana e Viana-Barbosa (2011)

Para a abordagem quantitativa, o trabalho apresenta tabelas com análises detalhadas sobre cada ponto analisado (relações, ramificações, hierarquia e links cruzados), conforme apresenta a Tabela 2.

Dessa forma o estudo conclui afirmando que é possível realizar avaliações cognitivas do aluno, considerando análises feitas em Mapas Conceituais dada a comparação com o

Tabela 2 – Scores do Mapa Conceitual

MAPA CONCEITUAL	Nota	CATEGORIAS		VÁLIDAS	CATEGORIAS		NÃO VÁLIDAS (-)	Pontos Validos
O que é Movimento?	3,07	Relações:	0	9	Relações	0	0	9
			3			0		
		Ramificações	3	15	Ramificações	0	0	15
			0			0		
			0			0		
			0			0		
		Lig. Cruzadas	1	10	Lig. Cruzadas	0	0	10
Exemplos	0		Níveis H.	1		5		
Total de Pontos no Mapa								39

Fonte: Moraes, Santana e Viana-Barbosa (2011)

mapa de referência do professor. O estudo levou em consideração ainda a evolução na utilização de mapas conceituais pelos alunos.

Partindo da análise desses artigos, evidencia-se que a Aprendizagem Significativa continua sendo matéria atual e com espaço para outros trabalhos e novas abordagens. Seguindo essa direção, nossa pesquisa utiliza essas referências para dar um embasamento teórico maior à nosso estudo, sinalizando a relevância em se utilizar as estratégias participativas em sala de aula, a ideia da exploração dos subsunçores, os conceitos de diferenciação progressiva e reconciliação integrativa e da classificação da taxonomia topológica, associando toda essa sistematização ao nível de aprendizagem adquirido, podendo assim relacionar esses temas com a Aprendizagem Significativa.

3.2 MAPAS CONCEITUAIS

Organizamos essa seção da seguinte forma: Construção de Mapas Conceituais segundo Aguiar e Correia (2013), Qualidade de Conceitos e Estrutura do Mapa Conceitual baseado nos estudos de Ruiz-Moreno *et al.* (2007) e Kowata, Cury e Boeres (2012), Critérios de Avaliação de Mapas Conceituais, para esse tema mencionamos Gowin e Alvarez (2005) e Carvajal, Cañas *et al.* (2006), Classificação de Mapas Conceituais segundo Cañas *et al.* (2006), Valerio, Leake e Cañas (2008) e Lovati *et al.* (2017), é importante anotar que esses dois últimos grupos se complementam para uma avaliação mais completa. Por fim, Mapas Conceituais e PLN citando os trabalhos de Pérez e Vieira (2005) e Gaspar *et al.* (2018).

Aguiar e Correia (2013) sintetizaram em seu estudo 4 parâmetros de referência para treinar usuários iniciantes, unido a teoria com a prática, são eles: clareza semântica

das proposições, pergunta focal, organização hierárquica dos conceitos e revisões contínuas. Com base nesses parâmetros os autores desenvolveram as atividades de treinamento que auxiliaram os alunos na aquisição de habilidades em trabalhos usando Mapas Conceituais. A Tabela de Clareza proposicional é o primeiro item de treinamento e pode ser visualizado na Tabela 3, essa abordagem demanda uma análise do conteúdo do mapa individualmente por parte do aluno.

Tabela 3 – Exemplo de tabela de Clareza Proposicional

Conceito Inicial	Termo de Ligação	Conceito Final	É possível entender essa mensagem?
a) Mapas conceituais	-	Proposições	Não
b) Mapas conceituais	para as	Proposições	Não
c) Mapas conceituais	são formados por	Proposições	Sim
d) Mapas conceituais	foram formados por	Proposições	Sim
e) Mapas conceituais	serão formados por	Proposições	Sim
f) Proposições	são formados por	Mapas conceituais	Sim
g) Mapas conceituais	não são formadas por	Proposições	Sim
h) Mapas conceituais	podem ser formados por	Proposições	Sim

Fonte: Aguiar e Correia (2013)

O Mapa Conceitual Semiestruturado é o segundo item do estudo e propõe incitar o aluno ao exercício de síntese, selecionando os conceitos mais significativos. Essa seleção não restringe o número de proposições que o aluno poderá elaborar, sempre vinculada a uma estrutura hierárquica bem definida. E, por último, a Atividade Colaborativa Expandida que envolve a revisão dos mapas em pares de forma colaborativa. Os resultados desse estudo demonstraram que essa metodologia traz proficiência para os alunos na criação de Mapas Conceituais.

Esse trabalho, evidencia a necessidade de se elaborar um material complementar com o objetivo de incentivar as boas práticas na construção de Mapas Conceituais. Por outro lado, há de se pensar na qualidade das proposições e na estrutura hierárquica dos mapas, para isso cita-se o estudo de Ruiz-Moreno *et al.* (2007), no qual argumenta que os Mapas Conceituais são ferramentas importantes para acompanhamento da aprendizagem, nesse sentido desenvolveram um trabalho que visa debater critérios de análise de mapas. Para a pesquisa, foram elencados critérios de avaliação embasados na quantidade e qualidade dos conceitos, os níveis de hierarquia, quantidade de ligações cruzadas, verbos de ligação, contextualização das proposições e estrutura do Mapa Conceitual.

Para o estudo, foram analisados 31 mapas divididos em duas áreas de conhecimento (17 mapas - Processo de Ensino-Aprendizagem e 14 mapas - Educação em Saúde). Com isso

pode-se considerar as seguintes contribuições: 80% dos mapas relacionados a “Processo de Ensino-Aprendizagem” possuem em média seis conceitos específicos em forma de exemplos, subordinados aos conceitos mais hierárquicos. Os mapas apresentam uma hierarquia média de 7 nós, as palavras de ligação entre os conceitos se mostraram, em sua maioria, úteis para inter-relacionar os conceitos. A avaliação da estrutura do Mapa Conceitual foi quantificada em sua maioria com estrutura de rede (60%), constando também relações cruzadas.

Kowata, Cury e Boeres (2012) discorreu em seu trabalho sobre a construção automática de Maps Conceituais por intermédio de textos em língua portuguesa do Brasil, possibilitando a redução das dificuldades de criação de mapas do “zero”. Após toda a parte de normalização dos textos, o protótipo computacional denominado Text2Cmap se mostrou uma solução viável para a criação de Mapas Conceituais baseados em textos em língua portuguesa, possibilitando a continuidade dos atributos semânticos dos documentos submetidos ao processo de concepção dos Mapas Conceituais. Com esses resultados, consegue-se embasamento necessário para implementar a normalização das proposições, etapa que será detalhada mais a frente.

Afim de identificar aspectos mais complexos, estudamos o *Scoring Criteria for Concept Maps (SCCM)* como métrica de avaliação complementar aos outros julgamentos aqui mencionados. Respalado pelo trabalho de Gowin e Alvarez (2005) onde leva em consideração toda a estrutura sintática, semântica e estrutural do mapeamento, o referido autor utiliza o protocolo SCCM para validar as proposições resultantes dos Mapas Conceituais.

Ressalta-se que para cada critério é atribuído uma pontuação específica, na qual todas essas variantes devem ser pontuadas quando válidas e quando não válidas, as não válidas estão em desconformidade com o contexto da pergunta focal.

- Hierarquia: Para cada nível de hierarquia é atribuído 5 (cinco) pontos;
- Ramificações: Na ocorrência de ramificações é atribuída pontuação específica a depender do nível onde ocorra a ramificação, sendo no Nível 1 (5 pontos), no Nível 2 (4 pontos), no Nível 3 (3 pontos), no Nível 4 (2 pontos) e do Nível 5 em diante (1 ponto);
- Relacionamentos: Para cada proposição válida teremos a atribuição de 3 pontos. Nesse critério, as ligações cruzadas não são definidas como proposições válidas, por isso não devem ser pontuadas;
- *Links* Cruzados: 10 pontos são associados a ocorrência de link cruzado. As ligações cruzadas indicam pensamento, capacidade criativa e consciência Gowin e Alvarez (2005).

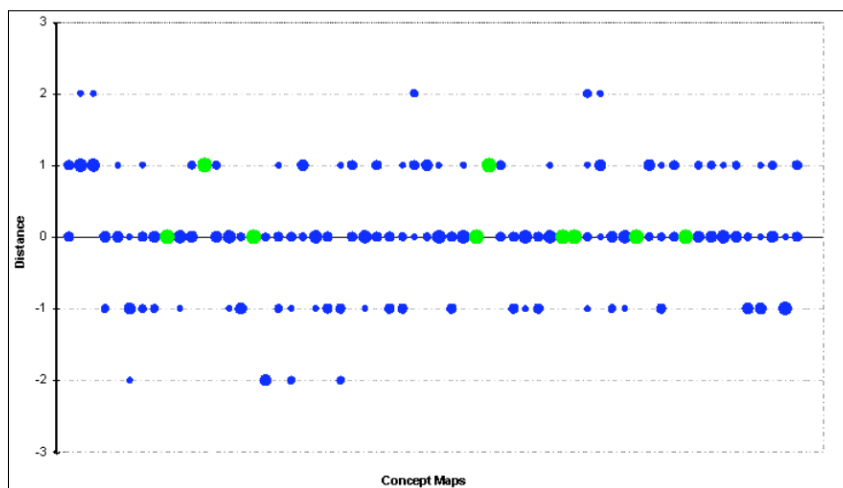
Seguindo a sequência, o trabalho de Carvajal, Cañas *et al.* (2006) propõe um estudo onde seu objetivo era avaliar Mapas Conceituais na visão de especialistas e não especialistas, o experimento tinha o intuito de classificar a qualidade do Mapa Conceitual visto por um período muito curto de tempo, tendo como referência a primeira impressão.

O avaliador julgava critérios como layout, inclusão da questão de foco, número de conceitos e a renderização do mapa. Para esse estudo, foram analisados 60 mapas e classificados como *muito bom*, *bom*, *ruim* ou *muito ruim*, desse quantitativo 13 mapas foram desenhados manualmente, 47 foram criados utilizando a ferramenta *CmapTools* Cañas *et al.* (2004), 12 mapas continham ícones em seus conceitos, 5 mapas foram desenhados e também criados por computador, 1 mapa foi exibido duas vezes, 2 mapa estavam em preto e branco, 3 mapas com o mesmo foram exibidos em layout diferentes, 18 mapas tiveram a questão focal modificada para nível de conceito e 9 mapas estavam com colocações diferentes nos conceitos.

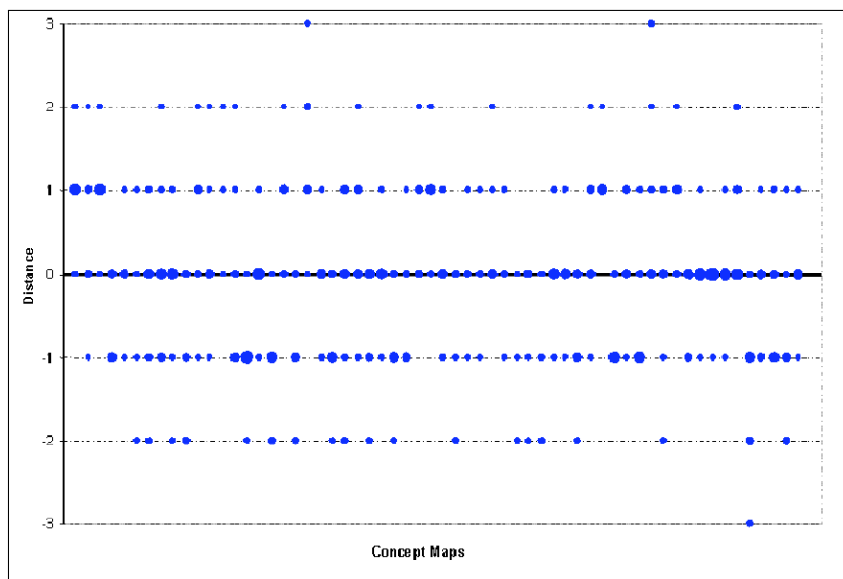
Os avaliadores foram divididos de acordo com a experiência em criação de Mapas Conceituais, 6 eram experientes e 19 não, porém há de se ressaltar que os 19 não especialistas já trabalhavam com Mapas Conceituais há vários anos.

Utilizando a classificação de Novak e Gowin (1984) para ranquear as avaliações, percebeu-se que os resultados geraram boas expectativas nos autores, a Figura 12 revela que há pouca variação dos especialistas em relação à avaliação da Novak, onde as bolas verdes foram os mapas classificados com a mesma pontuação do referido autor, e quanto mais próxima a pontuação menor é a distância para as bolas verdes. Já os não-especialistas Figura 13 tiveram um rendimento bem menor. Os resultados indicam que os especialistas tendem a concordar com a teoria de classificação de Novak com mais frequência do que os não-especialistas.

Figura 12 – Distância no *ranking* de classificação entre Novak e os Especialistas



Fonte: Carvajal, Cañas *et al.* (2006)

Figura 13 – Distância no *ranking* de classificação entre Novak e os Não-Especialistas

Fonte: Carvajal, Cañas *et al.* (2006)

Com o resultado desses experimentos, consegue-se o referencial teórico necessário para implementar a classificação dos Mapas Conceituais. Além disso utilizamos também Cañas *et al.* (2006), Valerio, Leake e Cañas (2008) e Lovati *et al.* (2017) para complementar o estudo.

Cañas *et al.* (2006) realizou um grande estudo sobre Mapas Conceituais nas escolas públicas do Panamá (*Projeto Conectar ao Conhecimento*) com o objetivo de qualificar esses mapas seguindo uma taxonomia pré-definida baseada em uma topologia que pondera a complexidade estrutural do mapa, sem levar em conta a relevância dos conceitos e proposições, focando apenas em medir o progresso do aluno iniciante, no momento de assimilação das questões fundamentais na construção do mapa.

Essa metodologia, conhecida como taxonomia topológica, consiste em 7 níveis de classificação, indo de 0 a 6. Para o estudo, foi considerado os seguintes quesitos: uso de conceitos em vez de partes do texto, as relações entre os conceitos, o grau de ramificação, hierárquica e a utilização de ligações cruzadas entre conceitos de “galhos” diferentes. Para o experimento fora utilizados 210 mapas, e o grau de concordância entre os avaliadores desses mapas foi definido pela porcentagem de concordância observada e coeficiente de *kappa*. Ao estabelecer essa metodologia, um mapa somente poderá pertencer a um determinado nível se e somente se contiver todas as características desse.

Dessa forma o classificador topológico inicia a análise pelo Nível 0 (nó raiz), caso não contenha proposições vinculadas a esse nó o mapa será classificado com Nível atual 0, caso contrário a verificação prossegue em $N(a) + 1$ até o Nível 6 ser alcançado. No decorrer da trajetória se houver falha na validação, a classificação será realizada em $N(a)$

- 1. Em um cenário de aprendizes, recomenda-se a implantação do Nível -1 para etiquetar tais mapas, caso apareçam.

Os níveis da taxonomia topológica são organizados conforme Tabela 4. No mesmo sentido Valerio, Leake e Cañas (2008) desenvolveram um estudo que utilizou a mesma taxonomia para validar mapas de forma automática, com isso os autores ressaltam que esse trabalho não tenta atribuir uma pontuação absoluta, mas apenas auxiliar o professor na análise quantitativa da estrutura dos mapas. O resultado desse experimento constatou que maioria dos mapas desenvolvidos pelos especialistas são classificados como de alto nível, segundo a taxonomia, confirmando o que foi dito em trabalhos anteriores em nosso estudo.

Tabela 4 – Condições necessárias para classificar a taxonomia topológica

Level #	Conditions by level as indicated in Cañas, Novak et al. (2006)	Conditions evaluated by to classify concept map M , which has concept c and linking phrase l the as nodes of M .
Level -1	No conditions	(default)
Level 0	At least 4 connected concepts Mostly long concept labels Empty linking-phrases 0 to 1 branching points	$\{ c : concept\} \geq 4$ (default) (default) (default)
Level 1	More concepts than long concept labels Half or more missing linking-phrases 0 to 1 branching points	$\{ c : concept \mid labelSize(c) < 12\} \geq \{ c : concept\} / 2$ (default) (default)
Level 2	More concepts than long concept labels Less than half missing linking-phrases At least 2 branching points	(checked at Level 1) $\{ l : linkingPhrase \mid labelSize(l) > 0\} \geq \{ l : linkingPhrase\} / 2$ $\mid branchingPoints(M) \geq 2$
Level 3	No long concept labels No linking-phrases missing At least 3 branching points Less than 3 hierarchy levels	$\forall c : concept, labelSize(c) < 12$ $\forall l : linkingPhrase, labelSize(l) > 0$ $\mid branchingPoints(M) \geq 3$ (default)
Level 4	No long concept labels No linking-phrases missing At least 5 branching points At least 3 hierarchy levels	(checked at Level 3) (checked at Level 3) $\mid branchingPoints(M) \geq 5$ $\exists c : concept, depth(c, M) > 3$
Level 5	No long concept labels No linking-phrases missing At least 5 branching points At least 3 hierarchy levels At least 1 cross-link	(checked at Level 3) (checked at Level 3) (checked at Level 4) (checked at Level 4) $\mid crossLinks(M) \geq 1$
Level 6	No long concept labels No linking-phrases missing At least 7 branching points At least 3 hierarchy levels At least 3 cross-links	(checked at Level 3) (checked at Level 3) $\mid branchingPoints(M) \geq 7$ (checked at Level 4) $\mid crossLinks(M) \geq 3$

Conditions labeled with (default) are met if the map fails a condition of a following level.
Conditions labeled with (checked at Level N) are revised at a previous level.

Fonte: Valerio, Leake e Cañas (2008)

A análise dos resultados da pesquisa indicou que a taxonomia topológica desenvolvida possui um nível de confiabilidade de moderada a boa, essa afirmação baseia-se no percentual de concordância entre os analisadores (55,7%) e, também, pelo coeficiente kappa simples (0,4778) com ponderação linear (0,7215).

Seguindo nesse cenário de trabalhos colaborativos, Lovati *et al.* (2017) propôs uma ferramenta de clusterização de Mapas Conceituais com o propósito de assistir o professor no momento de agrupar alunos com níveis cognitivos semelhantes. Com a clusterização, o autor conseguiu encontrar padrões de similaridade entres os mapas que possibilitaria o agrupamento dos aprendizes conforme retrata o modelo conceitual visto na Figura 14.

Figura 14 – Modelo Conceitual da Fermenta de Clusterização de Mapas Conceituais



Fonte: Lovati *et al.* (2017)

Os resultados do estudo apontaram que 90% dos mapas foram classificados de forma correta, fazendo com que a ferramenta consiga classificar os Mapas Conceituais em diferentes níveis cognitivos.

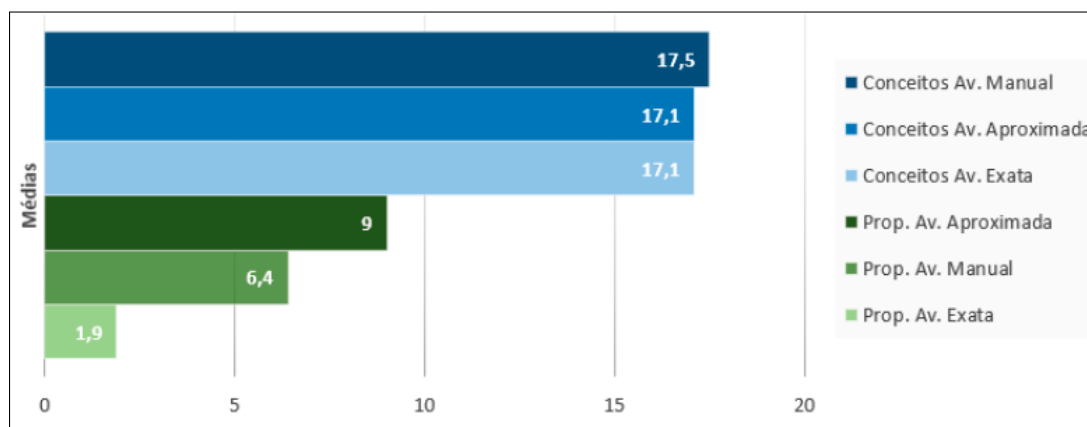
Além disso, se fez necessário a compreensão textual das proposições e, para essa lacuna utilizamos o Processamento de Linguagem Natural como apoio a essa compreensão. Grande parte dos artigos relacionados converte textos de diversas fontes em Mapas Conceituais, Pérez e Vieira (2005) utilizou essa técnica para desenvolver um protótipo que realiza a extração semiautomática de textos jornalísticos em mapas. Os referidos autores utilizaram ferramentas de PLN e o *parser* PALAVRAS¹ para realizar a parte de tokenização, análise sintática e processamento léxico-morfológico. Para validação foi realizada a comparação entre os mapas gerados pela fermenta que utilizavam textos completos e mapas gerados manualmente por dois voluntários utilizando resumos desses textos, os resultados indicaram baixa similaridade, fenômeno explicado pela diferença entre os formatos dos textos.

Já Gaspar *et al.* (2018) desenvolveu, em sua pesquisa, um protótipo apoiado em PLN que visa auxiliar o professor na utilização de Mapas Conceituais em sala de aula. O estudo aborda a construção de mapas a partir de textos, a comparação de mapas e a criação de mapas de domínio. Essa proposta utiliza metodologias bem embasadas na literatura e com carga de eficiência elevada. A utilização de mapas de referência propicia um ponto de controle na comparação e na direção do Mapa Conceitual do aluno. Na

¹ PALAVRAS é um parsers do projeto VISL (*Visual Interactive Syntax Learning*), do *Institute of Language and Communication* da University of Southern Denmark. O parser tem como entrada um conjunto de sentenças de um corpus, e retorna como saída a análise sintática das sentenças

Figura 15 visualizamos uma proximidade sensível entre a avaliação dos conceitos feita pelo professor comparada com a ferramenta. A avaliação das proposições teve uma diferença expressiva devido aos mapas dos alunos possuírem verbos diferentes.

Figura 15 – Comparação do Professor versus Comparação Automática



Fonte: Gaspar *et al.* (2018)

Para o autor, a comparação automática entre mapas apresentada no trabalho mostrou resultados satisfatórios, revelando que o protótipo é confiável e representa um espelho do nível da turma.

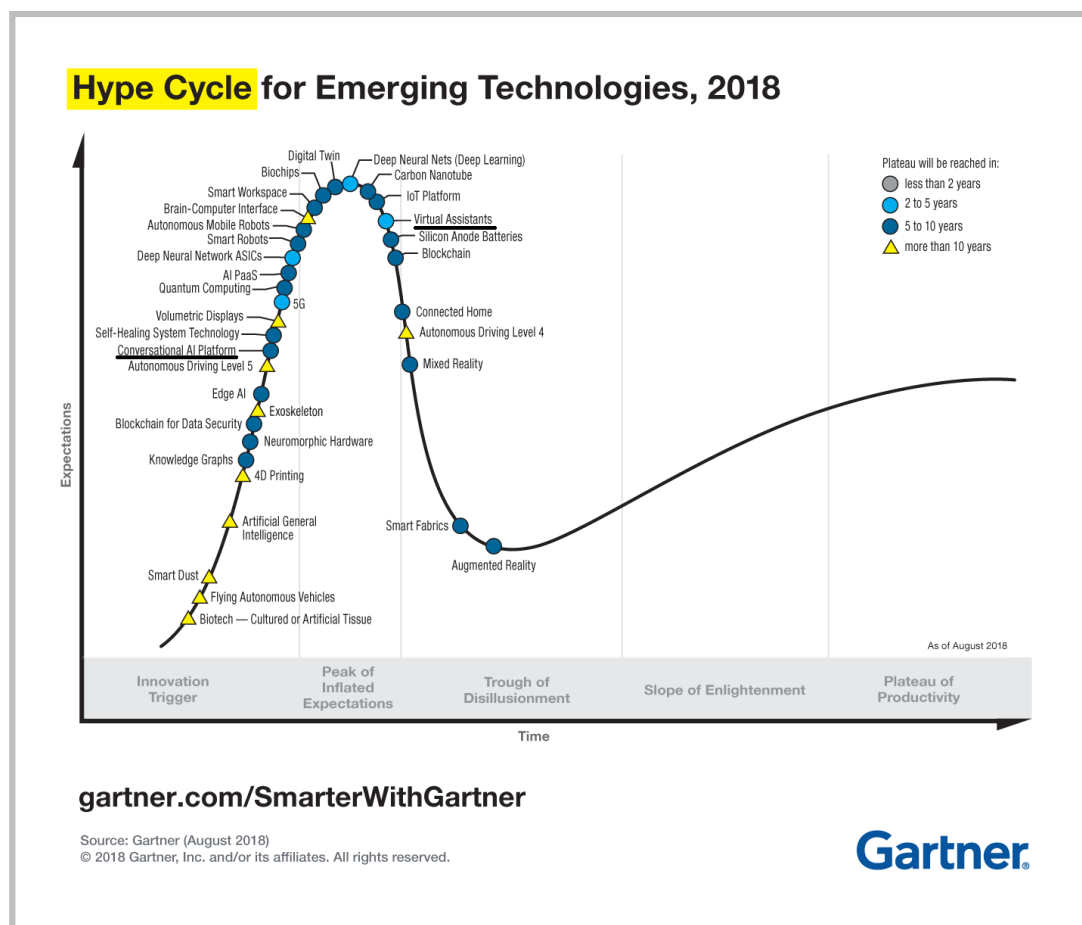
3.3 PROCESSAMENTO DE LINGUAGEM NATURAL

Diversos métodos têm sido implementados com o propósito de facilitar a avaliação da aprendizagem por meio de Mapas Conceituais, oportunizando, com isso, a aplicação de estratégias computacionais usando PLN. Considerando que essa área da Inteligência Artificial está em evidência conforme cita Gomes (2019), e promove meios diretos para a utilização de técnicas de processamento de grandes volumes de texto, ao mesmo tempo que consegue classificar essas informações, esses motivos justificam a importância dada a esse subitem.

Em nossa pesquisa estruturamos três grupos específicos para onde apontamos nossos esforços, são eles: Relevância do Tema, onde citamos Gartner (2018) e Gomes (2019), Processamento de Texto segundo Sette e Martins (2016) e Classificação de Texto conforme menciona os trabalhos de Loper e Bird (2002), Barbosa *et al.* (2017) e Môro *et al.* (2018).

Dessa forma, iniciamos esta seção objetivando investigar novos padrões que sustentem a utilização de PLN em nossa Arquitetura. De mais a mais, na Figura 16 percebemos essa tendência evidenciada pelo Gartner (2018) em seu anuário de tecnologias emergentes, mencionando assistentes virtuais e plataformas de conversação (em destaque).

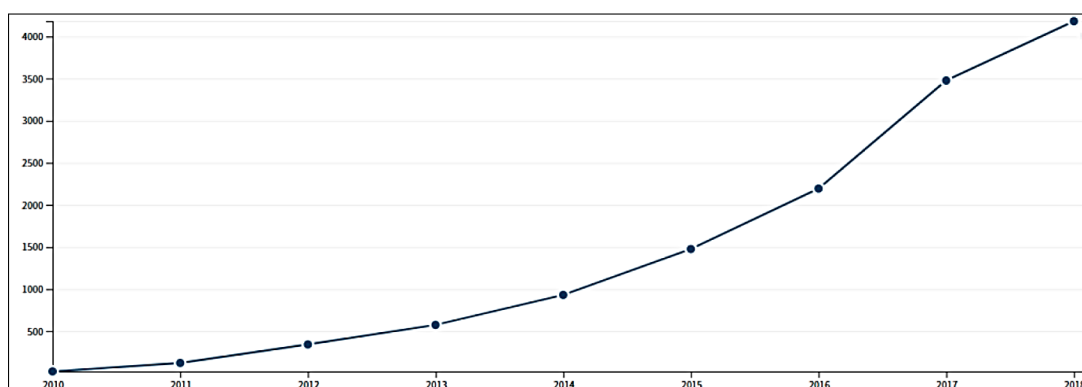
Figura 16 – Tecnologias Emergentes em 2018



Fonte: Adaptado de Gartner (2018)

Isso mostra que há um crescente interesse nessa área de conhecimento. Além disso, Gomes (2019) em seu estudo realizou levantamento sobre as buscas na base *Web of Science* relativas ao tema "processamento de linguagem natural", esse levantamento objetivou entender o cenário atual do tema em um contexto mais amplo, em outras palavras, buscou por "processamento de linguagem natural" em domínios como: direito, linguística, biomédica, ciência da computação e engenharias, conforme apresenta a Figura 17.

Figura 17 – Citações sobre o tópico “*natural language processing*” por ano



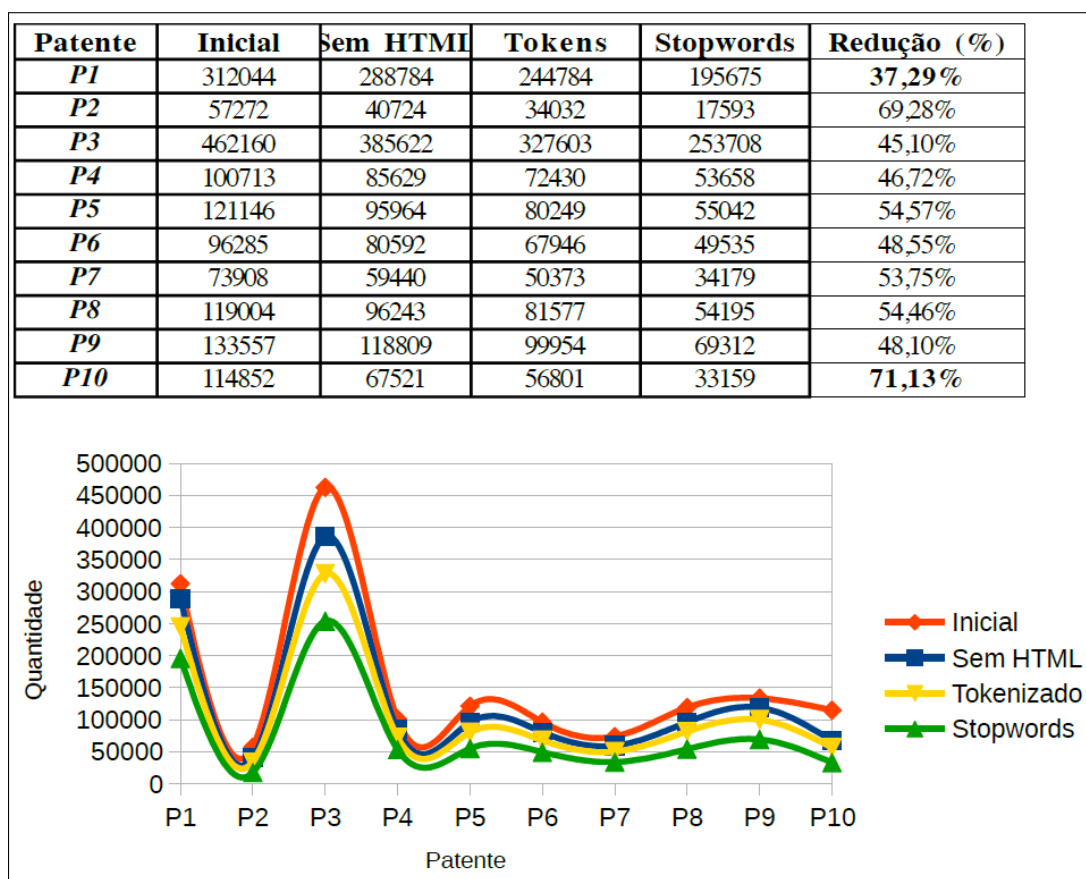
Fonte: Gomes (2019)

Sette e Martins (2016) encontraram uma lacuna na literatura relacionada a base de dados de patentes, com isso buscou em sua pesquisa desenvolver uma metodologia que coletasse, processasse, analisasse os dados e encontrasse padrões em documentos relacionados a patentes.

Esse estudo trabalhou fortemente no pré-processamento de texto, tema bastante discutido em nossa pesquisa, na Figura 18 visualizamos a evolução do pré-processamento de 10 patentes após cada técnica de remoção de caracteres não prioritários como: remoção de *tags*, *tokenização* e remoção de *stopwords*, com resultado notório na redução de palavras após o pré-processamento.

Essas técnicas de normalização de textos vão ao encontro da utilização de PLN em proposições extraídas de Mapas Conceituais, evidenciando assim a necessidade de seu uso para aprimorar os algoritmos classificação textuais.

Figura 18 – Comparação de quantidades de caracteres



Fonte: Sette e Martins (2016)

Com relação a classificação de textos, Loper e Bird (2002) discorrem sobre a documentação do NLTK em uma proposta mais teórica, abordando análise de parsing, classificação de textos, *tokenização*, regras de *chunk* e probabilidade. Em trabalho semelhante Barbosa *et al.* (2017) destacou a análise de sentimento, área de grande interesse na

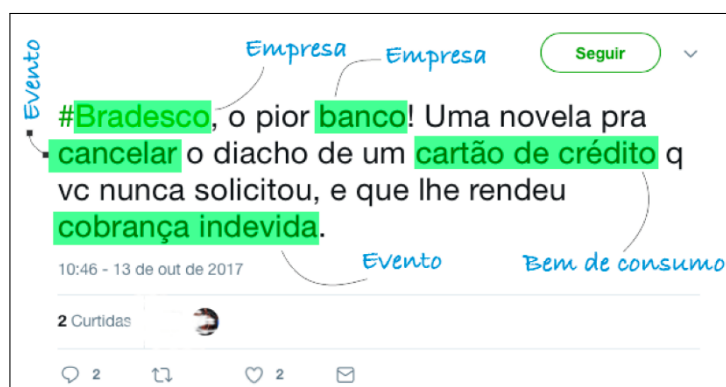
mineração de opinião.

Ambos os trabalhos abordam métodos específicos de classificar os textos, tema já tratado em nossa pesquisa. É oportuno mencionar que Joachims (2002) recomenda para a categorização de texto o algoritmo de *Naive Bayes*, retornando a probabilidade da proposição A ocorrer, dado o evento em B, levando em consideração a frequência em que os *tokens* da proposição são apresentados.

Do mesmo modo Môro *et al.* (2018) preocupou-se com a análise computacional dos documentos hospedados na internet em língua portuguesa, em especial o Reconhecimento de Entidades Nomeadas – REN que reconhece e classifica pessoas, lugares, organizações ou valores monetários.

Na Figura 19 temos um trecho de um tweet onde destacamos algumas Entidades Nomeadas classificadas, esse destaque é conhecido como anotação, onde sua classificação é chamada de entidade.

Figura 19 – Twitter com destaque para EN



Fonte: Môro *et al.* (2018)

O trabalho ainda traz referencial teórico relacionado a técnicas de pré-processamento de textos com auxílio de PLN, contudo adiciona aos resultados as métricas de avaliação das REN's referenciadas em spaCy e NLTK dentre outras.

A Tabela 5 mostra uma análise comparativa dos trabalhos mais relevantes que embasaram essa pesquisa, resgatando os resultados mais significativos ou que trazem teorias basilares (Aprendizagem Significativa, Mapas Conceituais e Processamento de Linguagem Natural).

Tabela 5 – Comparação dos Trabalhos Relacionados

Trabalhos	Aprendizagem Significativa	Mapas Conceituais	Processamento de Linguagem Natural
Loper e Bird (2002)	-	-	X
Jesus e Silva (2004)	X	X	-
Pérez e Vieira (2005)	-	X	X
Gowin e Alvarez (2005)	-	X	-
Carvajal, Cañas <i>et al.</i> (2006)	-	X	-
Moreira (2006)	X	-	-
Cañas <i>et al.</i> (2006)	-	X	-
Ruiz-Moreno <i>et al.</i> (2007)	-	X	-
Valerio, Leake e Cañas (2008)	-	X	-
Moraes; <i>et al.</i> , (2011)	X	X	X
Martins, Linhares e Reis (2011)	X	X	-
Braathen (2012)	X	-	-
Kowata, Cury e Boeres (2012)	-	X	-
Lemos (2013)	X	-	-
Aguiar e Correia (2013)	-	X	-
Sette e Martins (2016)	-	-	X
Lovati <i>et al.</i> (2017)	-	X	X
Barbosa <i>et al.</i> (2017)	-	-	X
Filho e Ferreira (2018)	X	-	-
Gaspar <i>et al.</i> (2018)	-	X	X
Môro <i>et al.</i> (2018)	-	-	X
Gartner (2018)	-	-	X
Gomes (2019)	-	-	X

Fonte: Autoria Própria

Por fim, os trabalhos relacionados tem o objetivo de embasar a pesquisa e encontrar lacunas na literatura. Em adição, a nossa pesquisa buscou agrupar estudos que fornecessem subsidio para classificação de Mapas Conceituais. Encontramos diversas pesquisas nesse sentido, algumas buscam apenas avaliar a estrutura do mapa, outra desenvolver metodologias de comparação. Entretanto, não encontramos trabalhos que agrupem avaliação estrutural, taxonomia e análise de objeto através de recursos computacionais envolvendo PLN.

4 Desenvolvimento da Arquitetura Computacional de Classificação de Mapas Conceituais

A solução desenvolvida nessa pesquisa representa um conjunto de métodos que classifica os Mapas Conceituais auxiliando na avaliação da cognição do aluno, essa avaliação é fundamentada nas proposições do Mapa Conceitual criado por esse mesmo aluno. O sistema estrutura uma base de dados específica baseado em um conteúdo proposto previamente, em seguida as proposições elaboradas na ferramenta *CmapTools* são exportadas para os módulos da nossa arquitetura, que de forma computacional estrutura, normaliza, valida, compara, nivela e por fim pontua essas proposições, etapas essas identificadas na Modelagem dos Processos apresentada na Figura 26. Então, toda essa sequência de passos retorna uma score que auxiliará a identificar o nível de aprendizado adquirido pelo aluno, respaldado na qualidade e na organização dos conceitos impressos no mapa.

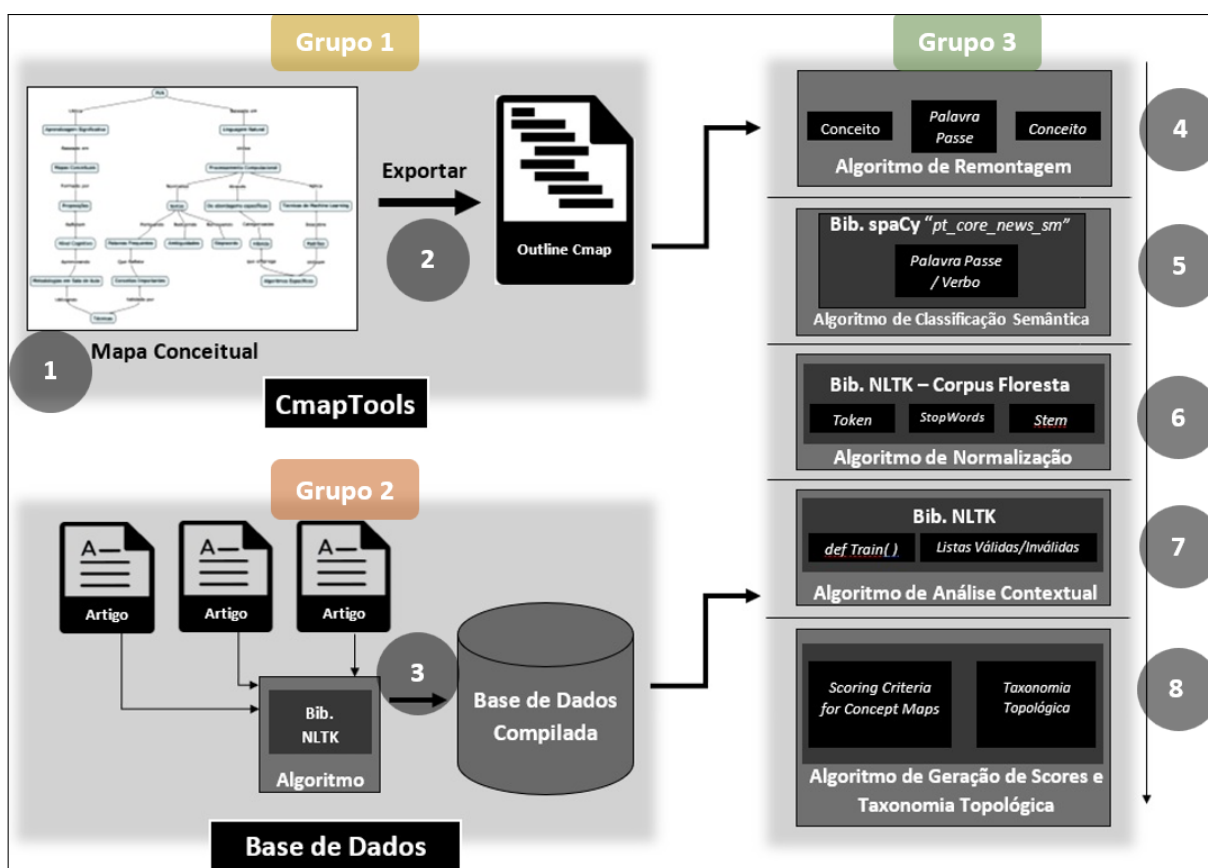
4.1 VISÃO GERAL

Com o propósito de apresentar nossa arquitetura, detalharemos todos os processos contidos nesse projeto organizado da seguinte forma: Iniciamos introduzindo a seção 4.2 – A Arquitetura, em seguida iniciamos o processo de validação da Arquitetura com a seção 4.3, logo após vem o subseção 4.3.1 - Criação dos Mapas Conceituais, em seguida 4.3.2 - Exportação do Arquivo Cmap, posterior a essa etapa iniciamos as argumentações da subseção 4.3.3 - Base de Dados, mais adiante começaremos a tratar a subseção 4.3.4 - Geração dos Arquivos de Entrada, o próximo passo será realizado na subseção 4.3.5 - Análise Geral das Proposições utilizando o algoritmo de remontagem, em seguida iniciaremos o processo da subseção 4.3.6 - Normalização e Segmentação de Texto, é nesse momento que realizamos a *tokenização*, a remoção dos *stopwords* e usamos a função de *stemming*, o processo seguinte ocorre na subseção 4.3.7 - Análise Contextual, nesse momento acontece a interação com a Base de Dados e as proposições são validadas, e por fim, as etapas 4.3.8 - Geração dos Scores e 4.3.9 - Geração do Nível da Taxonomia Topológica são executadas para concluir todo o processo. Ao final, convalidamos a pesquisa com a seção 4.4 - Resultados para expor as conclusões da nossa validação.

4.2 A ARQUITETURA

A modelagem da arquitetura foi dividida em três grupos conforme ilustrado na Figura 20. O primeiro grupo representa a fase inicial, remete a criação do Mapa Conceitual através do CmapTools e posterior exportação do arquivo no formato *Outline cmap*. O segundo grupo retrata a estruturação da base de dados, resultado da compilação dos textos do material de apoio, é oportuno explicar que a base de dados só iniciará sua interação com o terceiro componente no momento em que o processo do algoritmo de análise contextual for iniciado. O terceiro grupo trata da avaliação das proposições. Detalharemos a seguir todo o processo que compõe a arquitetura e como ocorre a comunicação entre os grupos para que a aplicação funcione de forma adequada.

Figura 20 – Arquitetura do Sistema



Fonte: Autoria Própria

O *Grupo 1* é composto pelo módulo “CmapTools” e responde pela criação do Mapa Conceitual (passo 1) e exportação das proposições para o formato Outline cmap (passo 2), etapa inicial que deve ser executado para se gerar o arquivo de saída. O desenvolvimento do mapa utiliza a ferramenta CmapTools que por sua vez já possui a função de exportar o mapa para o formato compatível com nossa arquitetura. É importante explicar que o *Grupo 1* é composto por ferramentas de terceiros, e se trata da única parte da arquitetura que não utilizou implementação extra.

Em seguida, no *Grupo 2*, há a estruturação da base de dados, nessa etapa o módulo “Base de Dados” reúne o material de apoio selecionado pelo especialista que será compilando em um arquivo único e ranqueado, gerando assim critérios para a validação das proposições (passo 3). Após esse estágio, o módulo “Base de Dados” fica em stand-by aguardando a interação com o *Grupo 3*.

Por fim, o *Grupo 3* é responsável pela validação e classificação das proposições, esse módulo contém 5 interações e inicia com o recebimento do arquivo *Outline cmap* (passo 4). A recepção do arquivo é realizada pelo Algoritmo de Remontagem que irá organizar as informações do arquivo em um modelo específico. Essa “remontagem” planifica o arquivo para a próxima etapa (passo 5) que realizará validação semântica das informações e prepara as proposições para a fase seguinte. A Normalização (passo 6) segmenta as proposições rotulando e adicionando atributos no *tokens*, nesse ponto as proposições ganham identidade permitindo assim a sua validação junto a base de dados.

Em seguida, o módulo “Banco de Dados” interage com as proposições, nesse momento o Algoritmo de Análise Contextual (passo 7) realiza inferência no banco de dados e identifica se a proposição está de acordo com as métricas propostas.

E, por último, temos a Geração dos Scores e Classificação Topológica do Mapa Conceitual (passo 8), responsável pela pontuação das proposições válidas e inválidas e posterior classificação topológica. É mister lembra que toda a arquitetura será detalhada em na próxima seção.

4.3 VALIDAÇÃO

A validação da arquitetura se faz verificando passo-a-passo cada elemento que a constitui, nas subseções a seguir iremos esmiuçar o funcionamento de cada módulo evidenciando seu funcionamento e realizado as validações com dados reais.

4.3.1 Criação Dos Mapas Conceituais

Perin, Cury e Menezes (2012) constatam em seu trabalho a versatilidade da utilização de Mapas Conceituais - MC, afirmando que é possível usar essa ferramenta para trabalhar o conteúdo de apenas uma única aula, uma unidade de um curso, um capítulo de um livro ou ainda para um programa educacional completo. Partindo desse princípio, iniciamos as explicações da nossa arquitetura escolhendo a turma, com 9 alunos, do Programa de Educação Tutorial PET/UERN do curso de Graduação em Ciência da Computação para elaboração dos mapas utilizados nesse projeto.

De maneira introdutória, elaboramos uma aula para apresentar aos alunos os conceitos de MC's, partindo de uma visão holística e indo até as técnicas de escolha e

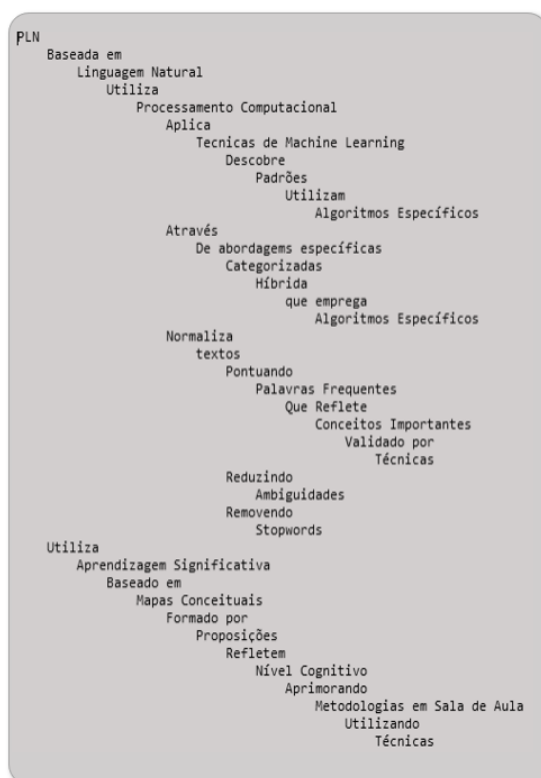
estruturação das proposições. Esse momento serviu para que os aprendizes pudessem tirar suas dúvidas sobre a metodologia e também de como utilizar o CmapTools para construir seus mapas, é importante ressaltar que o *CmapTools* implementa uma arquitetura flexível com instalação personalizada. Ao final da aula, todo o material foi disponibilizado para os alunos, inclusive uma videoaula explicando a instalação e utilização da ferramenta *CmapTools* em sua versão 6.03.

Além disso, o tema “Processamento de Linguagem Natural” foi escolhido para servir de base de conhecimento na criação dos mapas, três artigos científicos foram escolhidos para montar o referencial teórico, e com isso auxiliar os alunos no exercício. Ao final da atividade, os alunos foram instruídos a desenvolver o Mapa Conceitual nos moldes da pesquisa, conforme instruções repassadas a cada aluno e contidas no Apêndice A.

4.3.2 Exportação do Arquivo Cmap

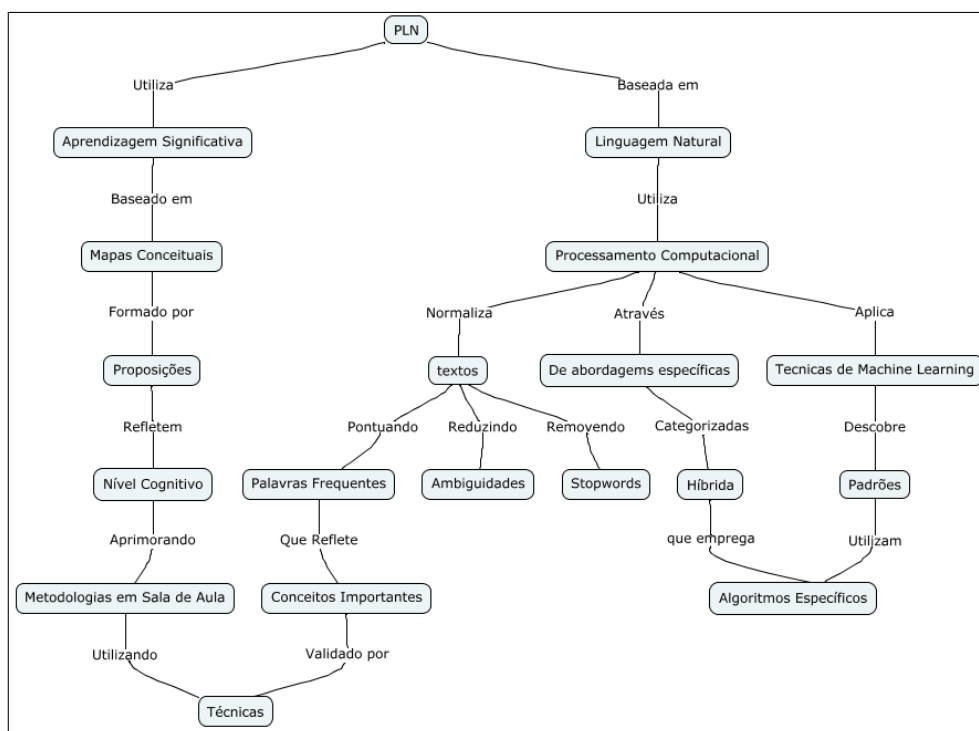
Uma característica importante do *CmapTools* é a possibilidade de exportar os mapas em diversos formatos, possibilitando que a utilização das informações contidas no mapa seja utilizada por outras ferramentas. Em nosso trabalho usamos o formato “*Outline do cmap...*” em estrutura de tópicos visto na Figura 21, esse formato traz um arquivo de saída indentado, semelhante a visualização de uma árvore. Na Figura 22 retrata o Mapa Conceitual utilizado como base para gerar a saída já mostrada.

Figura 21 – Saída do Mapa Conceitual Base



Fonte: Autoria Própria

Figura 22 – Mapa Conceitual Base



Fonte: Autoria Própria

4.3.3 Base de Dados

Essa fase é responsável pela criação e estruturação da base de dados, ela servirá de apoio para a busca de contextualidade das proposições extraídas dos Mapas Conceituais, em outras palavras, relacionaremos as proposições dos mapas com as informações contidas na base de dados e, através da mineração de informações, o algoritmo indicará se a proposição está de acordo com o conteúdo programático ou não.

Iniciamos esse subitem com a estruturação do conteúdo programático sobre Processamento de Linguagem Natural, compilamos os três artigos em um só arquivo para montar a base de dados “Base_NLP”, essa base é isenta de qualquer formatação, ou seja, trabalhamos com o texto livre que não possui nenhum “esquema” para descrever sua estrutura. Aditivamente criamos uma segunda base de dados para realizar a classificação comparativa do conteúdo das proposições, essa base foi construída com trabalhos relacionados a Sistemas de Informação “Base_SI”, salientamos, ainda, que a comparação mencionada se refere ao conteúdo das duas bases de dados.

Nesse projeto utilizamos mineração de dados para identificar e interpretar as informações impressas nas proposições. Em síntese, Madeira (2015) afirma que Mineração de Dados é um componente do processo de Descoberta de Conhecimento em Banco de Dados (*Knowledge Discovery in Databases – KDD*), esse processo engloba as tarefas de pré-processamento, voltado a remoção e normalização de dados, e as tarefas de

pós-processamento a que competem o reconhecimento de padrões e interpretação de modelos.

Para esse experimento adotamos uma abordagem probabilista visando credibilizar nossos resultados, essa abordagem foi referenciada no livro de Chakrabarti (2002) fundamentada no Algoritmo de *Naive Bayes*, um classificador probabilístico voltado para textos que utiliza probabilidade condicional, ou seja, infere a probabilidade do evento A ocorrer, dado o evento em B. Essa abordagem vem sendo muito utilizada hoje em dia na área de aprendizado de máquina. Para utilizar *Naive Bayes* precisamos transformar toda a base de dados em uma lista, essa lista foi dividida em frases, para realizar essa divisão utilizamos a Biblioteca NLTK do Python juntamente com o corpus Floresta¹ para identificar a estrutura das orações, sabendo assim onde encerra cada frase. Na Figura 23 exemplificamos como esse framework identifica as orações e distingue o “.” (ponto final) do “.” (ponto de abreviação) em destaque na figura, para isso utilizamos um texto sobre a vida do escritor Vinícius de Moraes.

Figura 23 – Texto com identificação de abreviações

['Marcus V. de Moraes nasceu em 1913 no bairro da Gávea, no Rio de Janeiro, filho de Clodoaldo Pereira da Silva Moraes, funcionário da Prefeitura, poeta e violinista amador, e Lygia Cruz de Moraes, pianista amadora.', 'Vinícius é o segundo de quatro filhos, Lygia (1911), Laetitia (1916) e Helius (1918).', 'Mudou-se com a família para o bairro de Botafogo em 1916, onde iniciou os seus estudos na Escola Primária Afrânio Peixoto.', 'Desde então, já demonstrava interesse em escrever poesias.', 'Em 1922, a sua mãe adoeceu e a família de Vinicius mudou-se para a Ilha do Governador, ele e sua irmã Lygia permanecendo com o avô, em Botafogo, para terminar o curso primário.', 'Marcus V. de Moraes ingressou em 1924 no Colégio Santo Inácio, de padres jesuítas, onde passou a cantar no coral e começou a montar pequenas peças de teatro.', 'Três anos mais tarde, tornou-se amigo dos irmãos Haroldo e Paulo Tapajós, com quem começou a fazer suas primeiras composições e a se apresentar em festas de amigos.', 'Em 1929, concluiu o ginásio e no ano seguinte, ingressou na Faculdade de Direito do Catete, hoje Faculdade Nacional de Direito (UFRJ).', 'Na chamada "Faculdade do Catete", conheceu e tornou-se amigo do romancista Otávio de Faria, que o incentivou na vocação literária.', 'Marcus V. de Moraes graduou-se em Ciências Jurídicas e Sociais em 1933.', 'Através de Otávio de Faria, começou a frequentar as reuniões da Ação Integralista Brasileira, movimento brasileiro de cunho nacionalista, na Livraria do Schmidt, chegando a se filiar ao mesmo.', 'Participou inclusive de algumas instruções das Milícias Integralistas no Arsenal de Guerra do Caju.', 'Em maio de 1937, publicou seu Soneto de Katherine Mansfield inédito na revista "Anauê!", nº 15.', 'Foi assinante de vários documentários Shell sobre o Integralismo.de carreira.]

Fonte: Autoria Própria

Após concluída essa etapa, identificamos cada oração com a respectiva classe do seu conteúdo, isso permitirá classificar e ranquear cada proposição. Na Figura 24 identificamos (em destaque) cada oração do conteúdo de NLP com sua respectiva classe, isso possibilitará que, após o ranqueamento dos conceitos, as proposições possam ser comparadas entre as classes NLP e SI.

¹ Floresta Sintáctica é um agrupamento de frases (*corpus*) analisadas morfológica e sintaticamente, e possui cerca de 6,7 milhões de palavras analisadas.

Figura 24 – Trecho da Base de Dados NLP classificada e dividida em frases

```
{'classe': 'nlp', 'frase': '3.'}, {'classe': 'nlp', 'frase': 'Extração de Termos Candidatos a
Conceitos A extração de termos pode ser feita segundo três abordagens: linguística, estatística
e híbrida.'}, {'classe': 'nlp', 'frase': 'A abordagem linguística faz uso de um nível
avançado de processamento linguístico, que geralmente envolve a identificação da organização
das frases dos textos (corpus) sintaticamente.'}, {'classe': 'nlp', 'frase': 'Para essa
análise utilizam-se programas de computador denominados parsers que fazem anotação sintática,
e ocasionalmente também semântica, de maneira automática.'}, {'classe': 'nlp', 'frase': 'A
partir da anotação feita pelo parser, outros programas utilizam essas anotações (geralmente
descritas em arquivos XML) como input adicional para a tarefa de identificação dos termos
candidatos a conceitos de uma ontologia.'}, {'classe': 'nlp', 'frase': 'Uma informação
linguística relevante para extrair automaticamente os termos de um corpus é a identificação
dos sintagmas nominais, isso permite a extração apenas de unidades de sentido previamente
identificadas pelo parser.'}, {'classe': 'nlp', 'frase': 'Um sintagma nominal é um conjunto de
palavras que exerce a mesma função gramatical que um substantivo.'}, {'classe': 'nlp',
'frase': 'Segundo (Kuramoto, 2002), ao contrário das palavras isoladas cujo significado
depende fortemente do contexto, sintagmas nominais tendem a permanecer com o mesmo significado
.'}, {'classe': 'nlp', 'frase': 'Na abordagem estatística os candidatos a termos são extraídos
segundo uma análise da sua frequência de ocorrência no corpus.'}, {'classe': 'nlp', 'frase':
'Existem ferramentas específicas que utilizam essa abordagem, por exemplo, o NSP - Ngrams
Statistic Package (Banerjee, 2003), que possui um conjunto de programas que realizam a tarefa
de extração de termos de um corpus de maneira totalmente estatística.'}, {'classe': 'nlp',
'frase': 'Esse método de Ngrams extrai termos composto por um número fixo de palavras.'},
```

Fonte: Autoria Própria

Cada proposição é dividida em *tokens* e os *tokens* são ranqueados, nessa classificação todos os sinais de pontuações são descartados para fins de validação. Na Figura 25 temos uma amostra do ranqueamento da base de dados “Base_NLP”, essa classificação leva em consideração a frequência em que a palavra é encontrada no conteúdo, Muller, Granatyr e Lessing (2015) afirmam que as sentenças mais importantes são aquelas em que os termos mais frequentes aparecem em maior quantidade dentro de uma mesma oração. Gonçalves (2012) fez um comparativo relevante sobre a frequência de palavras em textos, o autor afirma em seu estudo que “A presença do termo ‘Pelé’ aumenta em 5 vezes a chance de ocorrência dos termos ‘Copa’ e ‘1970’”.

Figura 25 – Trecho da Base de Dados Ranqueada

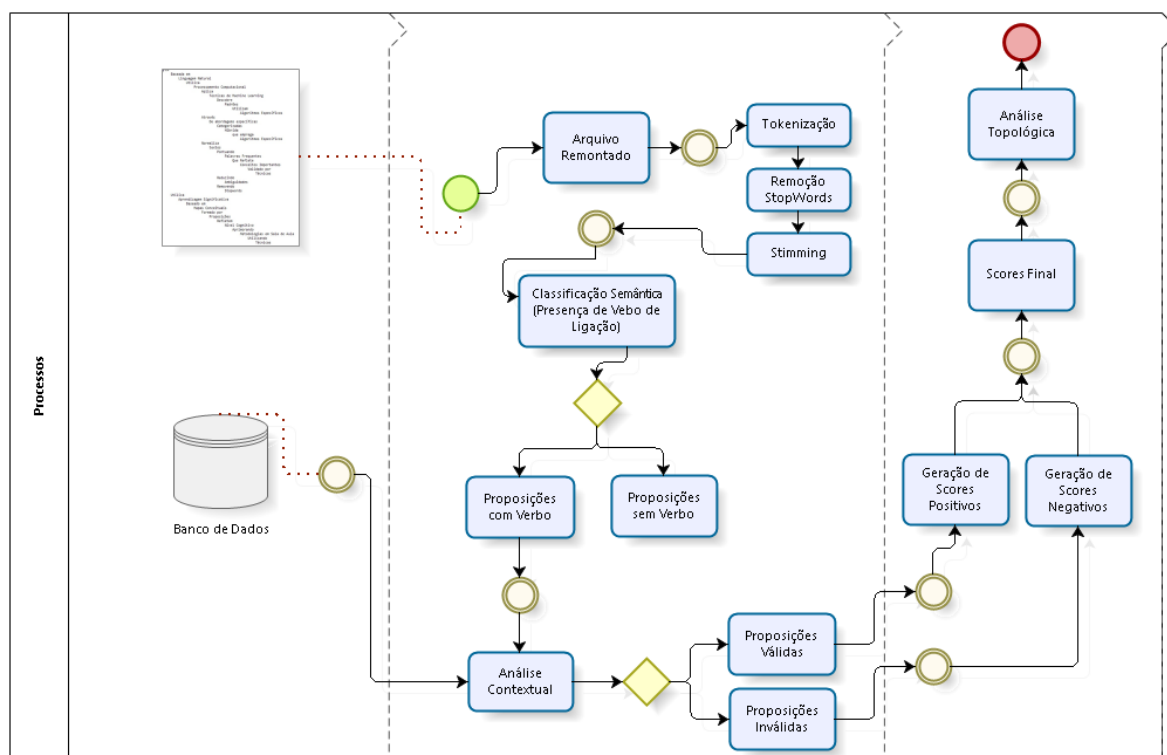
```
{'nlp': {'3': 2, '.': 163, 'extração': 17, 'termos': 18, 'candidatos': 3, 'conceitos': 4,
'pode': 8, 'ser': 14, 'feita': 4, 'segundo': 4, 'três': 2, 'abordagens': 5, ':': 64,
'linguística': 12, ',': 252, 'estatística': 7, 'híbrida': 2, 'abordagem': 12, 'faz': 3, 'uso':
13, 'nível': 2, 'avançado': 1, 'processamento': 17, 'linguístico': 1, 'geralmente': 3,
'envolve': 1, 'identificação': 5, 'organização': 1, 'frases': 1, 'textos': 12, '(': 50,
'corpus': 8, ')': 50, 'sintaticamente': 1, 'análise': 21, 'utilizam-se': 1, 'programas': 5,
'computador': 4, 'denominados': 1, 'parsers': 1, 'fazem': 3, 'anotação': 3, 'sintática': 6,
'ocasionalmente': 1, 'semântica': 7, 'maneira': 2, 'automática': 5, 'partir': 2, 'parser': 4,
'outros': 3, 'utilizam': 2, 'anotações': 1, 'descritas': 1, 'arquivos': 1, 'xml': 1, 'input': 1,
'adicional': 1, 'tarefa': 4, 'ontologia': 2, 'informação': 4, 'relevante': 4, 'extrair': 2,
'automaticamente': 1, 'é': 47, 'sintagmas': 5, 'nominais': 4, 'permite': 1, 'apenas': 3,
'unidades': 3, 'sentido': 5, 'previamente': 1, 'identificadas': 1, 'sintagma': 1, 'nominal':
2, 'conjunto': 3, 'palavras': 20, 'exerce': 1, 'mesma': 1, 'função': 2, 'gramatical': 2,
'substantivo': 2, 'kuramoto': 1, '2002': 1, 'contrário': 2, 'isoladas': 1, 'cujo': 1,
'significado': 4, 'depende': 1, 'fortemente': 1, 'contexto': 8, 'tendem': 1, 'permanecer': 1,
'extraídos': 4, 'frequência': 1, 'ocorrência': 1, 'existem': 2, 'ferramentas': 23,
'específicas': 1, 'exemplo': 13, 'nsp': 2, '-': 1, 'ngrams': 2, 'statistic': 1, 'package': 1,
```

Fonte: Autoria Própria

4.3.4 Geração dds Arquivos de Entrada

É conveniente evidenciar que o fluxo da informação a partir desse momento é mais complexo e merece uma atenção mais detalhada, para isso derivamos a Arquitetura do Sistema mostrada na Figura 20 com uma visão interna dos processos, pensando que dessa forma seja mais didática a apresentação de cada algoritmo. Na Figura 26 apresentamos os processos detalhados da nossa arquitetura, desde da exportação do arquivo de saída, passando pela estruturação da base de dados, continuando pelo arquivo remontado, em seguida passando pelo processo de normalização e classificação semântica, após essa etapa a interação com a base de dados e realizada e a análise contextual é efetivada e, por ultimo, realiza-se a geração dos *scores* e validação da taxonomia topológica.

Figura 26 – Modelagem dos Processos



Fonte: Autoria Própria

Ainda além, vimos a possibilidade de estruturar melhor o arquivo de saída exportado do *CmapTools*, adicionando características para melhorar a análise das proposições. A Figura 27 remonta as proposições, seguindo o formato: Nível | Conceito 1 | Verbo de Ligação | Conceito 2, deste modo conseguimos trabalhar cada proposição individualmente. Da mesma forma, também incluímos o nível onde a proposição se encontra no grafo, facilitando a identificação da diferenciação progressiva nas fases posteriores. É oportuno explicar que a estrutura de um Mapa Conceitual remeta ao conceito de grafos, onde seus conceitos se relacionam graficamente e os verbos de ligação representam as arestas, segundo

Lamas *et al.* (2005) a concepção gráfica de um MC pode ser definida como uma estrutura matemática denominada grafo.

Figura 27 – Arquivo de Saída “remontado”

```

N0 | PLN | Baseada | Linguagem Natural
N0 | PLN | Utiliza | Aprendizagem Significativa
N1 | Linguagem Natural | Utiliza | Processamento Computacional
N2 | Processamento Computacional | Aplica | Técnicas de Machine Learning
N2 | Processamento Computacional | Através | De abordagens específicas
N2 | Processamento Computacional | Normaliza | textos
N3 | Técnicas de Machine Learning | Descobre | Padrões
N4 | Padrões | Utilizam | Algoritmos Específicos
N3 | De abordagens específicas | Categorizadas | Híbrida
N4 | Híbrida | que emprega | Algoritmos Específicos
N3 | textos | Pontuando | Palavras Frequentes
N3 | textos | Reduzindo | Ambiguidades
N3 | textos | Removendo | Stopwords
N4 | Palavras Frequentes | Que Reflete | Conceitos Importantes
N5 | Conceitos Importantes | Validado por | Técnicas
N1 | Aprendizagem Significativa | Utiliza | Mapas Conceituais
N2 | Mapas Conceituais | Formado por | Proposições
N3 | Proposições | Refletem | Nível Cognitivo
N4 | Nível Cognitivo | Aprimorando | Metodologias em Sala de Aula
N5 | Metodologias em Sala de Aula | Baseada | Técnicas

```

Fonte: A autoria Própria

Os arquivos de entrada refletem o resultado da manipulação das informações que foram inseridas inicialmente no mapa. Com a saída do Mapa Conceitual Base o algoritmo percorre cada linha e armazena o conceito. Para realizar a travessia em toda a árvore, nossa arquitetura faz uso de técnicas de busca em profundidade, selecionando o primeiro conceito do arquivo de saída do Mapa Conceitual Base (Figura 21) como nó raiz, pois no momento da geração do arquivo o nó raiz é gravado na primeira linha. Após percorrer todos a árvore, é gerado o arquivo de saída “remontado” com a reorganização em linha de cada proposição do mapa.

4.3.5 Análise Geral das Proposições

Para análise das proposições Kowata, Cury e Boeres (2012) destacam abordagens diferentes a depender das características empregadas nos textos, para tanto existem os Métodos Linguísticos que se utilizam de técnicas de linguística computacional como análise morfológica, sintática, semântica, pragmática e do discurso. Já os Métodos Estatísticos são apoiados em indícios quantitativos, com mais ênfase em análise de frequência de repetição de termos ou referências de palavras, modelos estatísticos para definir peso que indica a relevância da palavra em um grupo de documentos e, por fim, técnicas voltadas à conjunto de documentos (clusterização). Por conseguinte, os Métodos Híbridos combinam técnicas

linguísticas, técnicas de aprendizagem de máquina e técnicas estatísticas, dessa forma não se consegue definir qual abordagem predomina.

Salienta-se ainda, que na realização desse trabalho utilizamos métodos híbridos para validar as proposições. Após a análise do arquivo de saída “remontado” mencionado na Figura 27, nosso algoritmo de validação identificava se a estrutura de cada proposição segue o padrão Conceito 1 | Verbo de Ligação | Conceito 2, essa validação reconhece apenas se o espaço está preenchido, para esse momento nenhuma análise mais apurada é realizada, caso o algoritmo encontre alguma proposição que não esteja de acordo com a estrutura mencionada, esse proposição é descartada da contagem dos scores. Na Figura 28 visualizamos a saída da análise supracitada, porém como todos os mapas analisados estavam com a estrutura correta tivemos que simular o erro na supressão da palavra passe (em destaque).

Figura 28 – Arquivo de Saída “remontado” sem a palavra passe

```

N0 | PLN | Baseada em | Linguagem Natural
N0 | PLN | Utiliza | Aprendizagem Significativa
N1 | Linguagem Natural | Utiliza | Processamento Computacional
N2 | Processamento Computacional | | De abordagens específicas
N2 | Processamento Computacional | Aplica | Tecnicas de Machine Learning
N2 | Processamento Computacional | Normaliza | textos
N3 | De abordagens específicas | Categorizadas | Híbrida
N4 | Híbrida | que emprega | Algoritmos Específicos
N3 | Tecnicas de Machine Learning | Descobre | Padrões
N4 | Padrões | Utilizam | Algoritmos Específicos
N3 | textos | Pontuando | Palavras Frequentes
N3 | textos | Reduzindo | Ambiguidades
N3 | textos | Removendo | Stopwords
N4 | Palavras Frequentes | Que Reflete | Conceitos Importantes
N5 | Conceitos Importantes | Validado por | Técnicas
N1 | Aprendizagem Significativa | Baseado em | Mapas Conceituais
N2 | Mapas Conceituais | Formado por | Proposições
N3 | Proposições | Refletem | Nível Cognitivo
N4 | Nível Cognitivo | Aprimorando | Metodologias em Sala de Aula
N5 | Metodologias em Sala de Aula | Utilizando | Técnicas
    
```

Fonte: Autoria Própria

Em seguida, as proposições selecionadas conforme norma anterior serão agrupadas do ponto de vista semântico. Essa validação leva em consideração se palavra passe é classificada semanticamente como verbo. A inexistência do termo de ligação inibe a compreensão da relação entre os conceitos, produzindo um mapa limitado com uma mensagem incompleta, onde sua estrutura não poderá ser classificada como proposição.

A forma de validação dos verbos descreveremos mais adiante, nos limitaremos nesse tópico a retratar as saídas tratadas pela arquitetura. Na Figura 29 visualizamos a saída do mapa produzido pelo aluno Alfa (nome sugestivo), nessa imagem são organizadas 27

proposições, da mesma forma na Figura 30 reclassificamos as proposições e extraímos apenas as que continham a palavra passe não classificada semanticamente como verbo.

Figura 29 – Arquivo de Saída do aluno Alfa

```
N0 | Processamento de linguagem natural | dedica-se | Desnvolver sistemas operacionais
N0 | Processamento de linguagem natural | Sub-área | Inteligência Artificial
N1 | Desnvolver sistemas operacionais | Entender | Através do pre-processamento
N1 | Desnvolver sistemas operacionais | Entender | Seres humanos
N2 | Através do pre-processamento | Verificando | Estrutura da língua
N3 | Estrutura da língua | Reduzindo | Vocabulário
N4 | Vocabulário | Utilizando | Correção ortográfica
N4 | Vocabulário | Utilizando | Normalização
N4 | Vocabulário | Utilizando | Remoção de numerais
N4 | Vocabulário | Utilizando | Remoção de stopwords
N4 | Vocabulário | Utilizando | Stemização e lematização
N5 | Remoção de stopwords | Remove | Palavras frequentes
N5 | Stemização e lematização | Reduz | Palavra ao radical
N5 | Stemização e lematização | Reduz | Palavra ao seu lema
N1 | Inteligência Artificial | Estuda | Capacidade da máquina
N2 | Capacidade da máquina | Entender | Através do pre-processamento
N2 | Capacidade da máquina | Entender | Seres humanos
N3 | Através do pre-processamento | Verificando | Estrutura da língua
N4 | Estrutura da língua | Reduzindo | Vocabulário
N5 | Vocabulário | Utilizando | Correção ortográfica
N5 | Vocabulário | Utilizando | Normalização
N5 | Vocabulário | Utilizando | Remoção de numerais
N5 | Vocabulário | Utilizando | Remoção de stopwords
N5 | Vocabulário | Utilizando | Stemização e lematização
N6 | Remoção de stopwords | Remove | Palavras frequentes
N6 | Stemização e lematização | Reduz | Palavra ao radical
N6 | Stemização e lematização | Reduz | Palavra ao seu lema
```

Fonte: Autoria Própria

Figura 30 – Classificação Semântica do aluno Alfa

```
----- ..:Proposições que não possuem verbo classificado semânticamente:. -----
['N0 | Processamento de linguagem natural | dedica-se | Desnvolver sistemas '
'operacionais\n',
'N0 | Processamento de linguagem natural | Sub-área | Inteligência '
'Artificial\n',
'N5 | Remoção de stopwords | Remove | Palavras frequentes\n',
'N6 | Remoção de stopwords | Remove | Palavras frequentes\n']
-----
```

Fonte: Autoria Própria

Para a classificação dos verbos de ligação, há a necessidade de se identificar as tags utilizadas pela biblioteca spaCy para mapear semanticamente as partes dos discursos, na Tabela 6 elencamos todas as *tags* utilizadas pela biblioteca spaCy para esse serviço. Ressaltamos, ainda, que para essa classificação a biblioteca utiliza o esquema de dependência universal que marca as principais categorias do discurso².

² Disponível em: <https://universaldependencies.org/u/pos/>

Tabela 6 – Tags da Biblioteca spaCy

Part of Speech - POS	Descrição	Exemplos
ADJ	adjetivo	grande, velho, verde, incompreensível, primeiro
ADP	aposição	em, para, durante
ADV	advérbio	muito, amanhã, para baixo, onde, lá
AUX	auxiliar	é, tem (feito), vai (faz), deve (faz)
CONJ	conjunção	e, ou, mas
CCONJ	conjunção coordenativa	e, ou, mas
DET	determinante	um, o
INTJ	interjeição	ai, bravo, olá
NOUN	substantivo	garota, gato, árvore, ar, beleza
NUM	numeral	1, 2017, um, setenta e sete, IV, MMXIV
PART	partícula	não
PRON	pronome	Eu, você, ele, ela, eu, eles mesmos, alguém
PROPN	nome próprio	Mary, John, Londres, OTAN, HBO
PUNCT	pontuação	., (,),?
SCONJ	conjunção subordinada	se, enquanto isso
SYM	símbolo	\$,%, §, ©, +, -, , , =, :)
VERB	verbo	correr, correr, correr, comer, comer, comer
X	de outros	sfpkdspxmsa
SPACE	espaço	

Fonte: Autoria Própria

Percebemos que na Figura 31 alguns verbos não foram classificados de forma correta, isso reflete o modelo “*pt_core_news_sm*”³ da biblioteca spaCy, esse pacote do Python usa modelos neurais para marcação, análise e reconhecimento de entidades, seu desempenho pode trazer certa deficiência em alguns gêneros como texto em mídias sociais, porém sua acurácia gira em torno de 88% e sua precisão refere-se às anotações "padrão prata" no corpus do WikiNER⁴.

³ Disponível em: https://github.com/explosion/spaCy-models/releases/tag/pt_core_news_sm-2.1.0

⁴ WikiNER são bases de dados formadas por artigos oriundo da Wikipedia disponíveis em nove idiomas, Alemão, Espanhol, Holandês, Polonês, Português, Inglês, Francês, Italiano e Russo.

Para compreender melhor a classificação dada a essas proposições, analisamos individualmente (em destaque) na Figura 31 cada classificação da palavra candidata a verbo para entender como se deu essa avaliação.

Figura 31 – Classificação Semântica errônea dos verbos do aluno Alfa

```
[('Processamento', 'NOUN'), ('de', 'ADP'), ('linguagem', 'NOUN'), ('natural', 'ADJ'), ('dedica-se', 'PROPN'), ('Desnvolver', 'PROPN'), ('sistemas', 'NOUN')]

[('Processamento', 'NOUN'), ('de', 'ADP'), ('linguagem', 'NOUN'), ('natural', 'ADJ'), ('Sub-área', 'NOUN'), ('Inteligência', 'PROPN')]

[('Remoção', 'NOUN'), ('de', 'ADP'), ('stopwords', 'NOUN'), ('Remove', 'PROPN'), ('Palavras', 'PROPN'), ('frequentes', 'ADJ')]

[('Remoção', 'NOUN'), ('de', 'ADP'), ('stopwords', 'NOUN'), ('Remove', 'PROPN'), ('Palavras', 'PROPN'), ('frequentes', 'ADJ')]
```

Fonte: Autoria Própria

Em contraponto, destacamos na Figura 32 as proposições que nosso algoritmo classificou como corretas, analisando esses destaques podemos constatar que 100% das classificações foram corretas, atingindo o máximo de acurácia. Para um entendimento mais apurado, incluímos todas as classificações realizadas pela nossa arquitetura no Apêndice B.

Figura 32 – Classificação Semântica correta dos verbos do aluno Alfa

```
[('através', 'ADV'), ('do', 'DET'), ('pre-processamento', 'NOUN'), ('verificando', 'VERB'), ('estrutura', 'NOUN'), ('da', 'ADP'), ('língua', 'NOUN')]

[('estrutura', 'NOUN'), ('da', 'ADP'), ('língua', 'NOUN'), ('reduzindo', 'VERB'), ('vocabulário', 'NOUN')]

[('vocabulário', 'NOUN'), ('utilizando', 'VERB'), ('correção', 'NOUN'), ('ortográfica', 'ADJ')]

[('vocabulário', 'NOUN'), ('utilizando', 'VERB'), ('normalização', 'NOUN')]

[('vocabulário', 'NOUN'), ('utilizando', 'VERB'), ('remoção', 'NOUN'), ('de', 'ADP'), ('numerais', 'NOUN')]

[('vocabulário', 'NOUN'), ('utilizando', 'VERB'), ('remoção', 'NOUN'), ('de', 'ADP'), ('stopwords', 'NOUN')]

[('vocabulário', 'NOUN'), ('utilizando', 'VERB'), ('stemização', 'NOUN'), ('e', 'CCONJ'), ('lematização', 'NOUN')]

[('stemização', 'NOUN'), ('e', 'CCONJ'), ('lematização', 'NOUN'), ('reduz', 'VERB'), ('palavra', 'NOUN'), ('a', 'ADP'), ('o', 'DET'), ('radical', 'NOUN')]

[('stemização', 'NOUN'), ('e', 'CCONJ'), ('lematização', 'NOUN'), ('reduz', 'VERB'), ('palavra', 'NOUN'), ('a', 'ADP'), ('o', 'DET'), ('seu', 'DET'), ('lema', 'NOUN')]

[('inteligência', 'NOUN'), ('artificial', 'ADJ'), ('estuda', 'VERB'), ('capacidade', 'NOUN'), ('da', 'ADP'), ('máquina', 'NOUN')]

[('capacidade', 'NOUN'), ('da', 'ADP'), ('máquina', 'NOUN'), ('entender', 'VERB'), ('através', 'ADV'), ('do', 'DET'), ('pre-processamento', 'NOUN')]

[('capacidade', 'NOUN'), ('da', 'ADP'), ('máquina', 'NOUN'), ('entender', 'VERB'), ('seres', 'NOUN'), ('humanos', 'ADJ')]

[('através', 'ADV'), ('do', 'DET'), ('pre-processamento', 'NOUN'), ('verificando', 'VERB'), ('estrutura', 'NOUN'), ('da', 'ADP'), ('língua', 'NOUN')]

[('estrutura', 'NOUN'), ('da', 'ADP'), ('língua', 'NOUN'), ('reduzindo', 'VERB'), ('vocabulário', 'NOUN')]

[('vocabulário', 'NOUN'), ('utilizando', 'VERB'), ('correção', 'NOUN'), ('ortográfica', 'ADJ')]

[('vocabulário', 'NOUN'), ('utilizando', 'VERB'), ('normalização', 'NOUN')]

[('vocabulário', 'NOUN'), ('utilizando', 'VERB'), ('remoção', 'NOUN'), ('de', 'ADP'), ('numerais', 'NOUN')]

[('vocabulário', 'NOUN'), ('utilizando', 'VERB'), ('remoção', 'NOUN'), ('de', 'ADP'), ('stopwords', 'NOUN')]

[('vocabulário', 'NOUN'), ('utilizando', 'VERB'), ('stemização', 'NOUN'), ('e', 'CCONJ'), ('lematização', 'NOUN')]

[('stemização', 'NOUN'), ('e', 'CCONJ'), ('lematização', 'NOUN'), ('reduz', 'VERB'), ('palavra', 'NOUN'), ('a', 'ADP'), ('o', 'DET'), ('radical', 'NOUN')]

[('stemização', 'NOUN'), ('e', 'CCONJ'), ('lematização', 'NOUN'), ('reduz', 'VERB'), ('palavra', 'NOUN'), ('a', 'ADP'), ('o', 'DET'), ('seu', 'DET'), ('lema', 'NOUN')]
```

Fonte: Autoria Própria

Após a validação semântica, buscou-se entender de forma contextual as proposições, e para isso havia a necessidade de comparar as proposições criadas pelos alunos com o conteúdo exposto no material de apoio.

Populamos a base de dados com o conteúdo de apoio disponibilizado para os alunos, esse material continha uma compilação de três textos acadêmicos sobre Processamento de Linguagem Natural, nesse contexto, comparamos as proposições validadas no passo anterior com as relações contextuais do banco de dados. No subitem 4.3.3 detalhamos sobre a montagem da base de dados e as comparações realizadas.

4.3.6 Normalização e Segmentação de Texto

Nesse momento, nossa solução importa o arquivo gerado no passo anterior e identifica cada um dos elementos que formam a proposição, esses elementos são chamados de *tokens*. Cada *token* representa um objeto específico, podendo ser uma palavra, um número e/ou sinais de pontuação, com isso o algoritmo reconhece e individualiza cada *token*, dessa forma há a possibilidade de trabalhar cada trecho e identificar a relevância do *token* para a proposição. Na Figura 33 *tokenizamos* todo o mapa do Aluno Alfa e transformamos em uma lista, com isso melhora-se o processamento e cria-se uma tabela de valores para as próximas etapas.

Figura 33 – *Tokenização* do mapa do Aluno Alfa

```
[ 'Processamento', 'de', 'linguagem', 'natural', 'dedica-se', 'Desnvolver', 'sistemas',
'operacionais', 'Processamento', 'de', 'linguagem', 'natural', 'Sub-área',
'Inteligência', 'Artificial', 'Desnvolver', 'sistemas', 'operacionais', 'Entender',
'Através', 'do', 'pre-processamento', 'Desnvolver', 'sistemas', 'operacionais',
'Entender', 'Seres', 'humanos', 'Através', 'do', 'pre-processamento', 'Verificando',
'Estrutura', 'da', 'língua', 'Estrutura', 'da', 'língua', 'Reduzindo', 'Vocabulário',
'Vocabulário', 'Utilizando', 'Correção', 'ortográfica', 'Vocabulário', 'Utilizando',
'Normalização', 'Vocabulário', 'Utilizando', 'Remoção', 'de', 'numerais',
'Vocabulário', 'Utilizando', 'Remoção', 'de', 'stopwords', 'Vocabulário', 'Utilizando',
'Stemização', 'e', 'lematização', 'Remoção', 'de', 'stopwords', 'Remove', 'Palavras',
'frequentes', 'Stemização', 'e', 'lematização', 'Reduz', 'Palavra', 'ao', 'radical',
'Stemização', 'e', 'lematização', 'Reduz', 'Palavra', 'ao', 'seu', 'lema',
'Inteligência', 'Artificial', 'Estuda', 'Capacidade', 'da', 'máquina', 'Capacidade',
'da', 'máquina', 'Entender', 'Através', 'do', 'pre-processamento', 'Capacidade', 'da',
'máquina', 'Entender', 'Seres', 'humanos', 'Através', 'do', 'pre-processamento',
'Verificando', 'Estrutura', 'da', 'língua', 'Estrutura', 'da', 'língua', 'Reduzindo',
'Vocabulário', 'Vocabulário', 'Utilizando', 'Correção', 'ortográfica', 'Vocabulário',
'Utilizando', 'Normalização', 'Vocabulário', 'Utilizando', 'Remoção', 'de', 'numerais',
'Vocabulário', 'Utilizando', 'Remoção', 'de', 'stopwords', 'Vocabulário',
'Utilizando', 'Stemização', 'e', 'lematização', 'Remoção', 'de', 'stopwords',
'Remove', 'Palavras', 'frequentes', 'Stemização', 'e', 'lematização', 'Reduz',
'Palavra', 'ao', 'radical', 'Stemização', 'e', 'lematização', 'Reduz', 'Palavra',
'ao', 'seu', 'lema']
```

Fonte: Autoria Própria

É oportuno salientar que é nessa fase onde as palavras consideradas desnecessárias são eliminadas, esses *tokens* não possuem relevância semântica no contexto em que existem, por isso são removidos na fase de pré-processamento. Na Figura 34, utilizamos a função

Stopwords do NLTK para remover os *tokens* irrelevantes e, na Figura 35, disponibilizamos a lista de *StopWords* do corpus Floresta utilizadas nesse experimento.

Figura 34 – Remoção dos *StopWords* do mapa do Aluno Alfa

```
[ 'Processamento', 'linguagem', 'natural', 'dedica-se', 'Desnvolver', 'sistemas',
'operacionais', 'Processamento', 'linguagem', 'natural', 'Sub-área', 'Inteligência',
'Artificial', 'Desnvolver', 'sistemas', 'operacionais', 'Entender', 'Através',
'pre-processamento', 'Desnvolver', 'sistemas', 'operacionais', 'Entender', 'Seres',
'humanos', 'Através', 'pre-processamento', 'Verificando', 'Estrutura', 'língua',
'Estrutura', 'língua', 'Reduzindo', 'Vocabulário', 'Vocabulário', 'Utilizando',
'Correção', 'ortográfica', 'Vocabulário', 'Utilizando', 'Normalização', 'Vocabulário',
'Utilizando', 'Remoção', 'numerais', 'Vocabulário', 'Utilizando', 'Remoção',
'stopwords', 'Vocabulário', 'Utilizando', 'Stemização', 'lematização', 'Remoção',
'stopwords', 'Remove', 'Palavras', 'frequentes', 'Stemização', 'lematização', 'Reduz',
'Palavra', 'radical', 'Stemização', 'lematização', 'Reduz', 'Palavra', 'lema',
'Inteligência', 'Artificial', 'Estuda', 'Capacidade', 'máquina', 'Capacidade',
'máquina', 'Entender', 'Através', 'pre-processamento', 'Capacidade', 'máquina',
'Entender', 'Seres', 'humanos', 'Através', 'pre-processamento', 'Verificando',
'Estrutura', 'língua', 'Estrutura', 'língua', 'Reduzindo', 'Vocabulário',
'Vocabulário', 'Utilizando', 'Correção', 'ortográfica', 'Vocabulário', 'Utilizando',
'Normalização', 'Vocabulário', 'Utilizando', 'Remoção', 'numerais', 'Vocabulário',
'Utilizando', 'Remoção', 'stopwords', 'Vocabulário', 'Utilizando', 'Stemização',
'lematização', 'Remoção', 'stopwords', 'Remove', 'Palavras', 'frequentes',
'Stemização', 'lematização', 'Reduz', 'Palavra', 'radical', 'Stemização',
'lematização', 'Reduz', 'Palavra', 'lema']
```

Fonte: Autoria Própria

Figura 35 – Lista dos *StopWords* do corpus Floresta

```
[ 'de', 'a', 'o', 'que', 'e', 'do', 'da', 'em', 'um', 'para', 'com', 'não', 'uma', 'os',
'no', 'se', 'na', 'por', 'mais', 'as', 'dos', 'como', 'mas', 'ao', 'ele', 'das', 'à',
'seu', 'sua', 'ou', 'quando', 'muito', 'nos', 'já', 'eu', 'também', 'só', 'pelo',
'pela', 'até', 'isso', 'ela', 'entre', 'depois', 'sem', 'mesmo', 'aos', 'seus',
'quem', 'nas', 'me', 'esse', 'eles', 'você', 'essa', 'num', 'nem', 'suas', 'meu',
'às', 'minha', 'numa', 'pelos', 'elas', 'qual', 'nós', 'lhe', 'deles', 'essas',
'esses', 'pelas', 'este', 'dele', 'tu', 'te', 'vocês', 'vos', 'lhes', 'meus',
'minhas', 'teu', 'tua', 'teus', 'tuas', 'nosso', 'nossa', 'nossos', 'nossas', 'dela',
'delas', 'esta', 'estes', 'estas', 'aquele', 'aquela', 'aqueles', 'aquelas', 'isto',
'aquilo', 'estou', 'está', 'estamos', 'estão', 'estive', 'esteve', 'estivemos',
'estiveram', 'estava', 'estávamos', 'estavam', 'estivera', 'estivéramos', 'esteja',
'estejamos', 'estejam', 'estivesse', 'estivéssemos', 'estivessem', 'estiver',
'estivermos', 'estiverem', 'hei', 'hã', 'havemos', 'hã', 'houve', 'houvemos',
'houveram', 'houvera', 'houvéramos', 'haja', 'hajamos', 'hajam', 'houvesse',
'houvéssemos', 'houvessem', 'houver', 'houvermos', 'houverem', 'houverei', 'houverá',
'houveremos', 'houverão', 'houveria', 'houveríamos', 'houveriam', 'sou', 'somos',
'são', 'era', 'éramos', 'eram', 'fui', 'foi', 'fomos', 'foram', 'fora', 'fôramos',
'seja', 'sejamos', 'sejam', 'fosse', 'fôssemos', 'fossem', 'for', 'formas', 'forem',
'serei', 'será', 'seremos', 'serão', 'seria', 'seríamos', 'seriam', 'tenho', 'tem',
'temos', 'tém', 'tinha', 'tínhamos', 'tinham', 'tive', 'teve', 'tivemos', 'tiveram',
'tivera', 'tivéramos', 'tenha', 'tenhamos', 'tenham', 'tivesse', 'tivéssemos',
'tivessem', 'tiver', 'tivermos', 'tiverem', 'terei', 'terá', 'teremos', 'terão',
'teria', 'teríamos', 'teriam']
```

Fonte: Autoria Própria

Outra etapa da normalização e segmentação de texto em nossa arquitetura é a utilização do algoritmo de *stemming*, esse código tem a função de reduzir as palavras a uma forma comum de representação denominada *stem* (radical), com isso os afixos oriundos de derivação ou de flexão são eliminados, possibilitando assim que conceitos morfológicamente correlatos sejam citados de forma comum, facilitando a sua combinação. Exemplificando a afirmação, temos as palavras “Utilizo”, “Utilizou” e “Utilizando”, derivadas do verbo

“Utilizar”, com a aplicação do algoritmo de *stemming* todas as palavras mencionadas anteriormente seriam reduzidas ao stem “Utiliz”, isso permite melhorar o índice de similaridade entre os conceitos obrigatórios. Na Figura 36 temos as proposições do Aluno Alfa com todos os seus conceitos reduzidos ao radical.

Figura 36 – Proposições do Aluno Alfa após aplicação do Algoritmo de *stemming*

```
[ 'process', 'lingu', 'natur', 'dedica-s', 'desnvolv', 'sistem', 'operac', 'process',
  'lingu', 'natur', 'sub-áre', 'intelig', 'artific', 'desnvolv', 'sistem', 'operac',
  'entend', 'através', 'pre-process', 'desnvolv', 'sistem', 'operac', 'entend', 'ser',
  'human', 'através', 'pre-process', 'verific', 'estrut', 'língu', 'estrut', 'língu',
  'reduz', 'vocabul', 'vocabul', 'utiliz', 'correç', 'ortográf', 'vocabul', 'utiliz',
  'normal', 'vocabul', 'utiliz', 'remoç', 'numer', 'vocabul', 'utiliz', 'remoç', 'stopword',
  'vocabul', 'utiliz', 'stemiz', 'lemat', 'remoç', 'stopword', 'remov', 'palavr', 'frequ',
  'stemiz', 'lemat', 'reduz', 'palavr', 'radic', 'stemiz', 'lemat', 'reduz', 'palavr', 'lem',
  'intelig', 'artific', 'estud', 'capac', 'máquin', 'capac', 'máquin', 'entend', 'através',
  'pre-process', 'capac', 'máquin', 'entend', 'ser', 'human', 'através', 'pre-process',
  'verific', 'estrut', 'língu', 'estrut', 'língu', 'reduz', 'vocabul', 'vocabul', 'utiliz',
  'correç', 'ortográf', 'vocabul', 'utiliz', 'normal', 'vocabul', 'utiliz', 'remoç', 'numer',
  'vocabul', 'utiliz', 'remoç', 'stopword', 'vocabul', 'utiliz', 'stemiz', 'lemat', 'remoç',
  'stopword', 'remov', 'palavr', 'frequ', 'stemiz', 'lemat', 'reduz', 'palavr', 'radic',
  'stemiz', 'lemat', 'reduz', 'palavr', 'lem']
```

Fonte: Autoria Própria

4.3.7 Análise Contextual

Essa fase é responsável por identificar se a proposição criada pelo aluno será classificada pelo algoritmo *Naive Bayes* como pertencente a “Base_NLP” ou “Base_SI”, os *tokens* presentes na proposição são minerados dentro do conteúdo das bases e é calculado a probabilidade da sequência de *tokens* que forma a proposição ocorrer entre as duas bases de dados disponibilizadas.

Ao realizar a avaliação, a arquitetura gera duas listas de saída, uma com as proposições validadas conforme algoritmo de *Naive Bayes* (Lista Válida) e uma segunda lista contendo as proposições reprovadas (Lista Inválida), conforme visualizamos na Figura 37 e Figura 38.

Para nossa pesquisa as proposições contidas na “Lista Inválida” referem-se as orações que não condizem com o contexto exposto no conteúdo programático. Em um segundo momento, o especialista convalida as mesmas proposições atestando a sua coerência, todas as avaliações estão contidas no Apêndice C.

Figura 37 – Lista de proposições válidas do Aluno Alfa

```

----- .:Correção de Contexto - Validas:. -----
['N2 | Através do pre-processamento | Verificando | Estrutura da língua\n',
'N3 | Estrutura da língua | Reduzindo | Vocabulário\n',
'N4 | Vocabulário | Utilizando | Correção ortográfica\n',
'N4 | Vocabulário | Utilizando | Normalização\n',
'N4 | Vocabulário | Utilizando | Remoção de numerais\n',
'N4 | Vocabulário | Utilizando | Remoção de stopwords\n',
'N4 | Vocabulário | Utilizando | Stemização e lematização\n',
'N5 | Stemização e lematização | Reduz | Palavra ao radical\n',
'N5 | Stemização e lematização | Reduz | Palavra ao seu lema\n',
'N1 | Inteligência Artificial | Estuda | Capacidade da máquina\n',
'N2 | Capacidade da máquina | Entender | Através do pre-processamento\n',
'N2 | Capacidade da máquina | Entender | Seres humanos\n',
'N3 | Através do pre-processamento | Verificando | Estrutura da língua\n',
'N4 | Estrutura da língua | Reduzindo | Vocabulário\n',
'N5 | Vocabulário | Utilizando | Correção ortográfica\n',
'N5 | Vocabulário | Utilizando | Normalização\n',
'N5 | Vocabulário | Utilizando | Remoção de numerais\n',
'N5 | Vocabulário | Utilizando | Remoção de stopwords\n',
'N5 | Vocabulário | Utilizando | Stemização e lematização\n',
'N6 | Stemização e lematização | Reduz | Palavra ao radical\n',
'N6 | Stemização e lematização | Reduz | Palavra ao seu lema\n']
-----

```

Fonte: Autoria Própria

Figura 38 – Lista de proposições inválidas do Aluno

```

----- .:Correção de Contexto - Invalidas:. -----
['N1 | Desnvolver sistemas operacionais | Entender | Através do '
'pre-processamento\n',
'N1 | Desnvolver sistemas operacionais | Entender | Seres humanos\n']
-----

```

Fonte: Autoria Própria

4.3.8 Geração dos Scores

De forma complementar, inserimos dois módulos de apoio à avaliação dos mapas. Nesse subitem, avaliamos quantitativamente as proposições validadas na fase anterior (Lista Válida e Lista Inválida) para a geração de um *score*, essa pontuação leva em conta características peculiares contidas nas proposições, auxiliando qualitativamente e classificando quantitativamente o Mapa Conceitual do discente, na Figura 39 ilustramos a validação das proposições do Aluno Alfa e no Apêndice D inserimos a relação contendo a pontuação dos 9 mapas avaliados. As características específicas utilizadas para pontuar os mapas estão presentes no Anexo A.

A geração dos *scores* permite inferir, com mais propriedade, a classificação cognitiva do aluno, pois essa sistemática abre caminho para uma avaliação profunda de todos os aspectos do Mapa Conceitual produzido pelo discente. Avalia a diferenciação progressiva firmada na profundidade das proposições válidas, atua ainda na reconciliação integrativa no momento que o crosslink é validado. Recomenda-se a utilização dos resultados dos *scores* juntamente com a Taxonomia Topológica, permitindo uma visão conceitual e estrutural macro do mapa Conceitual.

Figura 39 – *Scores* do Aluno Alfa

```

----- .:Scores para Proposições Válidas:. -----
21 proposicoes
relacionamentos * 3 = 63
maior hierarquia * 5 = 30
score de ramificacoes: 3
score positivo = 96

----- .:Scores para Proposições Inválidas:. -----
2 proposicoes
relacionamentos * 3 = 6
maior hierarquia * 5 = 5
Ramificações
score de ramificacoes: 0
score negativo = 11

----- CROSS -----

[]

----- CROSS -----

Score Total 85

```

Fonte: Autoria Própria

4.3.9 Geração do Nível da Taxonomia Topológica

O segundo módulo de avaliação inserido em nossa arquitetura foi a análise da estruturação do Mapa Conceitual, para esse evento foi considerada apenas a disposição dos conceitos, palavras de ligação, profundidade da árvore e reconexão com galhos adjacentes na estrutura hierárquica do mapa.

A Figura 40 mostra a classificação do mapa do Aluno Alfa segundo taxonomia proposta por Cañas *et al.* (2006), essa avaliação complementa o módulo anterior e possibilita ao professor ter uma visão macro do Mapa Conceitual, avaliando quantitativamente por meio dos *scores* e qualitativamente baseada na taxonomia topológica. No Apêndice E inserimos a relação contendo a classificação dos 9 mapas avaliados e no Anexo B consta o

protocolo da taxonomia topológica de Cañas *et al.* (2006).

Figura 40 – Classificação do mapa do Aluno Alfa

```
.: Sem Categoria :.  
.: Entre N4 e N5 :.  
Sem explicações longas: 24  
Não faltam Palavras de Link (Verbo de ligação): 24  
Ramificação Média: 8  
Níveis de Hierarquia: 6  
Ligações Cruzadas 0
```

Fonte: Aatoria Própria

4.4 RESULTADOS

Para geração dos resultados, validamos 197 proposições elaboradas pelos alunos nas seguintes classes: Estrutura Proposicional, Análise Semântica dos Verbos e Contexto, esse último validado com o conteúdo programático.

A validação ocorreu em três etapas, a primeira foi feita pelo especialista, proposição por proposição, onde o mesmo aprovava ou reprovava os trechos em cada classe, a segunda validação ocorreu com a análise automática da arquitetura proposta, também nas três classes, e por último foi comparado o resultado da etapa 1 com o da etapa 2, gerando uma acurácia de (100%) na estrutura da proposição, que leva em consideração o modelo CONCEITO 1 | VERBO | CONCEITO 2, DE (78%) da estrutura gramatical, que leva em consideração a presença do verbo classificado gramaticalmente e, por fim, (84%) de contexto, que avalia se a proposição está de acordo com o conteúdo programático. Essa análise é provada seguindo a métrica do teste de hipótese⁵. Com o experimento validado e os resultados obtidos, o trabalho se revela condizente com o que se propõe.

⁵ Teste de Hipóteses Estatística é uma metodologia que assessora a tomada de decisões e assegura escolher entre a hipótese nula H_0 ou hipótese alternativa H_a , baseando-se na amostra.

5 Considerações Finais

Para esta seção, reservamos às conclusões obtidas com a implementação da nossa pesquisa, do mesmo modo destinamos um momento para expor as limitações encontrada no decorrer do estudo e, de forma complementar, sugerimos os trabalhos que podem derivar desse trabalho.

5.1 CONCLUSÃO

Esse trabalho permitiu compreender a relação entre a Classificação de Maps Conceituais e a influência desses na Aprendizagem Significativa, dessa forma percebemos a real necessidade de utilizar uma arquitetura que pudesse mensurar esses mapas e categorizar essa avaliação para retratar a compreensão do aluno sobre um determinado tema.

Para se alcançar esse entendimento, definimos no escopo da pesquisa sete objetivos específicos. O primeiro discutiu a Teoria da Aprendizagem Significativa, buscando referenciar nosso trabalho com os conceitos necessários para se alcançar a aprendizagem de fato. Relacionamos diversos trabalhos que justificaram a escolha dessa teoria, comprovando a aprendizagem de conceitos como verdadeiramente significativa, e assim possibilitando um aprendizado mais substancial. Os conceitos abordados foram a diferenciação progressiva e a reconciliação integrativa, para tanto desenvolvemos um algoritmo que conseguia analisar as proposições extraídas do mapa do aprendiz e identificava essas ideias através da profundidade do mapa e da sua ligação cruzada, conceitos esses essenciais para o segundo objetivo.

Em seguida tratou-se dos requisitos necessários à criação dos Mapas Conceituais, onde foi desenvolvido um material complementar (Apêndice A) para auxiliar os discentes na elaboração dos seus esquemas gráficos. O material abordou a parte vital na criação dos mapas, a estruturação do conhecimento, que deve ser moldada no padrão hierárquico. Analisou-se ainda a parte conceitual, que refletia a ideia da relação entre conceitos e a escolha desses para a criação de um mapa que refletisse o aprendizado do aluno. Com isso, tivemos como resultado a predominância de mapas classificados no Nível 2 e Nível 4, para a conferência detalhada dos resultados, recomenda-se a análise dos Apêndice E.

Concomitantemente se trabalhou o terceiro, quarto e quinto objetivos, que trabalhavam diretamente com as proposições. O terceiro objetivo realizava a extração das proposições exportadas do *CmapTools*, essa etapa preparava as proposições para serem classificadas na fase que chamamos de Normalização que ocorreu no quarto objetivo. Em nossa pesquisa processamos 9 mapas conceituais num total de 197 proposições, que equivale

a 1179 *tokens* e a remoção de 251 *stopwords*. Para mais detalhamento recomenda-se a leitura do Apêndice B.

O quinto objetivo foi a validação das proposições, para esse momento desenvolvemos o algoritmo de análise contextual que validou as proposições conforme a sua classificação na base de dados, após essa validação o especialista realizou igual procedimento, chegando a uma acuraria de 100% na validação da estrutura da proposição, que leva em consideração o esquema CONCEITO | VERBO | CONCEITO, 78% na análise gramatical, que observa se o verbo foi classificado corretamente, e por fim a validação contextual, que observa se a proposição está dentro do contexto do conteúdo programático, nesse quesito nosso algoritmo chegou a uma acuraria de 84%. Para mais detalhamento recomenda-se a leitura do Apêndice C.

Os três últimos objetivos referem-se à elaboração de um modelo classificador de Mapas Conceituais conforme grau de aquisição de conhecimento do aprendiz. Como resultado, tivemos o índice de classificação desses mapas com uma média de pontuação de 37,88. Essa validação foi realizada com a implementação da métrica de Gowin e Alvarez (2005). Para a conferência detalhada dos resultados, recomenda-se a análise do Apêndice D.

Em síntese, desenvolvemos uma arquitetura computacional que analisa as proposições com o objetivo de classificar Mapas Conceituais, relacionado essa classificação aos conceitos de Aprendizagem Significativa. Sob uma ótica computacional, consideramos que o estudo foi bem-sucedido, na medida em que especificou e validou cada processo da arquitetura. Sob um contexto mais amplo e em consonância com os exemplos elencados em nosso referencial, há de se considerar que as respostas aos assuntos principais da pesquisa foram atendidas, comprovando assim a sua viabilidade.

5.2 LIMITAÇÕES

Essa pesquisa apresenta as seguintes limitações:

- A Arquitetura não foi validada em um ambiente de sala de aula convencional, ou seja, os alunos não estavam matriculados em uma disciplina nem tiveram um tempo maior para a continuidade do conteúdo, apenas tiveram uma aula expositiva sobre o tema. Por esse motivo não houve como validar a ferramenta de forma mais apurada. Uma vez que a arquitetura possa ser trabalhada em um ambiente de sala de aula, caberia nova pesquisa para investigar as características dos mapas conceituais produzidos;
- A validação da proposições segundo o seu contexto só foi realizada apenas por um especialista, com isso o seu resultado poderia ser outro caso a validação pudesse ser realizada por mais especialistas, um novo cenário poderia surgir;

- A quantidade de mapas analisados representa um volume baixo, nem todos os alunos que participaram da aula expositiva desenvolveram o mapa, isso representa uma amostragem baixa e pode interferir na percepção da análise do grupo.

5.3 TRABALHOS FUTUROS

Apesar dos resultados relevantes alcançados nesse trabalho, identificou-se lacunas que admitem a continuidade e/ou produção de novos trabalhos. Destacam-se, como trabalhos futuros, os seguintes pontos:

- Validação: Realizar validação com um número maior de especialistas, com uma população maior o refinamento da arquitetura pode ser mais apurado;
- Conhecimento: Utilizar discentes com níveis de conhecimento sobre Mapas Conceituais mais apurados, dessa forma conseguiríamos aprimorar o mecanismo de *scores* e a taxonomia topológica;
- Base de Dados: A possibilidade de se trabalhar com uma base de dados mais abrangente, que utilize mecanismos de *webservice* e consiga introduzir rapidamente conteúdo de repositório específicos em bases de dados de contextos diferentes;
- Base de Dados: Outro ponto relevante seria a forma de seleção dos textos para compor a base de dados para a validação contextual das proposições. Para isso, recomenda-se a realização de experimentos que permitam a avaliação dos resultados com textos de uma ou outra fonte e/ou repositório;
- Ambiente Virtual de Aprendizagem: Introduzir essa arquitetura como módulo em um Ambiente Virtual de Aprendizagem - AVA, integrando as funcionalidades de gerenciamento de turmas e acompanhamento de discente, oportunizando um ambiente mais amigável para o professor;
- Comparação de Mapas: Integra função que compara os mapas da turma e classifica os discentes dentro desse contexto. Essa funcionalidade possibilita montar grupos de estudos com alunos de níveis diferentes e buscar o nivelamento da turma;
- Feedback: Implementar função que viabilize a correção das validações do sistema, com isso o professor poderá alterar a validação da proposição que foi classificada de forma errônea, além disso, implementar técnicas de *Machine Learning* que consiga entender a correção e aplicar para situações iguais no futuro;
- *Stemming*: O uso da função de *stemming* pode causar a compreensão errônea no contexto da proposição, recomenda-se testes mais aprofundados para resolver esse problema;

Referências

- AGUIAR, J. G. de; CORREIA, P. R. M. Como fazer bons mapas conceituais? estabelecendo parâmetros de referências e propondo atividades de treinamento. *Revista Brasileira de pesquisa em Educação em Ciências*, v. 13, n. 2, p. 141–157, 2013. Citado 4 vezes nas páginas 26, 30, 39 e 40.
- AUSUBEL, D. P. *The psychology of meaningful verbal learning*. Grune & Stratton, 1963. Citado na página 36.
- AUSUBEL, D. P.; NOVAK, J. D.; HANESIAN, H. *Psicologia educacional*. [S.l.]: Interamericana, 1980. Citado 4 vezes nas páginas 17, 21, 24 e 37.
- AUSUBEL, D. P. *et al.* *Educational psychology: A cognitive view*. holt, rinehart and Winston New York, 1968. Citado 3 vezes nas páginas 16, 24 e 36.
- BANDEIRA, M. *Antologia de poetas brasileiros bissextos contemporâneos*. [S.l.]: Organização Simões, 1964. Citado na página 32.
- BARBOSA, E. F.; MOURA, D. G. de. Metodologias ativas de aprendizagem na educação profissional e tecnológica. *Boletim Técnico do Senac*, v. 39, n. 2, p. 48–67, 2013. Citado na página 19.
- BARBOSA, J. L. N. *et al.* Introdução ao processamento de linguagem natural usando python. *III ESCOLA REGIONAL DE INFORMÁTICA DO PIAUÍ*, v. 1, p. 336–360, 2017. Citado 2 vezes nas páginas 46 e 48.
- BLOOM, B. S. Some major problems in educational measurement. *Journal or Educational Research*, The University of Chicago Press, v. 38, n. 1, p. 139–142, 1944. Citado na página 17.
- BLOOM, B. S. Innocence in education. *The School Review*, The University of Chicago Press, v. 80, n. 3, p. 333–352, 1972. Citado na página 17.
- BLOOM, B. S. *et al.* *Handbook on formative and summative evaluation of student learning*. ERIC, 1971. Citado na página 17.
- BRAATHEN, P. C. Aprendizagem mecânica e aprendizagem significativa no processo de ensino-aprendizagem de química. *Revista eixo*, v. 1, n. 1, p. 63–69, 2012. Citado 5 vezes nas páginas 23, 24, 25, 36 e 37.
- CAÑAS, A. J. *et al.* Cmaptools: A knowledge modeling and sharing environment. Universidad Pública de Navarra, 2004. Citado na página 42.
- CAÑAS, A. J. *et al.* Confiabilidad de una taxonomía topológica para mapas conceptuales. Universidad de Costa Rica, 2006. Citado 7 vezes nas páginas 19, 38, 39, 43, 68, 69 e 96.
- CARVAJAL, R.; CAÑAS, A. J. *et al.* Assessing concept maps: first impressions count. In: CITESEER. *San José, Costa Rica: Universidad de Costa Rica*. [S.l.], 2006. Citado 3 vezes nas páginas 39, 42 e 43.

CARVALHO, W. S. *Reconhecimento de entidades mencionadas em português utilizando aprendizado de máquina*. Tese (Doutorado) — Universidade de São Paulo, 2012. Citado 2 vezes nas páginas 32 e 33.

CHAKRABARTI, S. *Mining the Web: Discovering knowledge from hypertext data*. [S.l.]: Elsevier, 2002. Citado na página 56.

CICUTO, C. A. T.; CORREIA, P. R. M. *et al.* Análise de vizinhança: uma nova abordagem para avaliar a rede proposicional de mapas conceituais. *Revista Brasileira de Ensino de Física*, São Paulo, SciELO Brasil, v. 34, n. 1, p. 1401–1, 2012. Citado na página 28.

DORNELES, P. F.; ARAUJO, I. S.; VEIT, E. A. Simulação e modelagem computacionais no auxílio à aprendizagem significativa de conceitos básicos de eletricidade: Parte i—circuitos elétricos simples. *Revista Brasileira de Ensino de Física*, SciELO Brasil, v. 28, n. 4, p. 487–496, 2006. Citado na página 36.

FAVA, R. Educação 3.0: aplicando o pdca nas instituições de ensino. *São Paulo: Saraiva*, p. 256, 2014. Citado na página 16.

FILHO, O. L. da S.; FERREIRA, M. Teorias da aprendizagem e da educação como referenciais em práticas de ensino: Ausubel e lipman. *Revista do Professor de Física*, v. 2, n. 2, 2018. Citado na página 27.

FILHO, R. L. L. S. *et al.* A evasão no ensino superior brasileiro. *Cadernos de pesquisa*, SciELO Brasil, v. 37, n. 132, p. 641–659, 2007. Citado na página 17.

GALVES, C.; FARIA, P. Tycho brahe parsed corpus of historical portuguese. *URL: <http://www.tycho.iel.unicamp.br/tycho/corpus/en/index.html>*, 2010. Citado na página 33.

GARTNER. *Trends Emerge in the Gartner Hype Cycle for Emerging Technologies*. 2018. Disponível em: <<https://www.gartner.com/smarterwithgartner/5-trends-emerge-in-gartner-hype-cycle-for-emerging-technologies-2018/>>. Citado 2 vezes nas páginas 46 e 47.

GASPAR, W. *et al.* Uma arquitetura tecnológica para apoiar o professor na utilização de mapas conceituais em sala de aula. In: *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*. [S.l.: s.n.], 2018. v. 29, n. 1, p. 158. Citado 3 vezes nas páginas 39, 45 e 46.

GOMES, D. Processamento de linguagem natural em português e aprendizagem profunda para o domínio de óleo e gás. *arXiv preprint arXiv:1908.01674*, 2019. Citado 3 vezes nas páginas 31, 46 e 47.

GOMES, M. J. *et al.* Challenges 2017: aprender nas nuvens, learning in the clouds: atas da x conferência internacional de tecnologias de informação e comunicação na educação. In: UNIVERSIDADE DO MINHO. CENTRO DE COMPETÊNCIA TIC (CCTIC UM). *Challenges 2017: aprender nas nuvens, learning in the clouds: atas da X Conferência Internacional de Tecnologias de Informação e Comunicação na Educação*. [S.l.], 2018. p. 1–1942. Citado na página 18.

- GONÇALVES, E. C. Mineração de texto. *SQL Magazine*, n. 105, 2012. Citado na página 57.
- GOWIN, D. B. *Educating*. [S.l.]: Cornell University Press, 1981. Citado na página 36.
- GOWIN, D. B.; ALVAREZ, M. C. *The art of educating with V diagrams*. [S.l.]: Cambridge University Press, 2005. Citado 3 vezes nas páginas 39, 41 e 71.
- GOWIN, D. B.; NOVAK, J. D.; VALADARES, C. *Aprender a aprender*. [S.l.: s.n.], 1996. Citado na página 36.
- JESUS, M. d.; SILVA, R. C. O. A teoria de david ausubel—o uso dos organizadores prévios no ensino contextualizado de funções. *VII Encontro Nacional de Educação Matemática—2004*, 2004. Citado 2 vezes nas páginas 36 e 37.
- JOACHIMS, T. *Learning to classify text using support vector machines*. [S.l.]: Springer Science & Business Media, 2002. v. 668. Citado na página 49.
- JOHNSON-LAIRD, P. N. *Mental models: Towards a cognitive science of language, inference, and consciousness*. [S.l.]: Harvard University Press, 1983. Citado na página 36.
- KOWATA, J. H. *Uma Abordagem para Construção Automática de Mapas Conceituais a partir de textos em Língua Portuguesa (Brasil)*. Dissertação (Mestrado) — Universidade Federal do Espírito Santo, 2010. Citado 3 vezes nas páginas 16, 28 e 33.
- KOWATA, J. H.; CURY, D.; BOERES, M. C. S. Em direção à construção automática de mapas conceituais a partir de textos. *Revista Brasileira de Informática na Educação*, v. 20, n. 1, p. 33, 2012. Citado 4 vezes nas páginas 34, 39, 41 e 59.
- LAMAS, F. *et al.* Comparando mapas conceituais utilizando correspondência de grafos. In: *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*. [S.l.: s.n.], 2005. v. 1, n. 1, p. 24–27. Citado na página 59.
- LEMOES, E. dos S. A aprendizagem significativa: estratégias facilitadoras e avaliação. *Série-Estudos-Periódico do Programa de Pós-Graduação em Educação da UCDB*, n. 21, 2013. Citado 2 vezes nas páginas 36 e 37.
- LOPER, E.; BIRD, S. Nltk: the natural language toolkit. *arXiv preprint cs/0205028*, 2002. Citado 2 vezes nas páginas 46 e 48.
- LOVATI, M. *et al.* Clusterizando mapas conceituais para identificar desempenho cognitivo em grupos. In: *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*. [S.l.: s.n.], 2017. v. 28, n. 1, p. 1397. Citado 3 vezes nas páginas 39, 43 e 45.
- MADEIRA, R. d. O. C. *Aplicação de técnicas de mineração de texto na detecção de discrepâncias em documentos fiscais*. Tese (Doutorado), 2015. Citado na página 55.
- MAGRINI, D. F. *et al.* Organização das nações unidas para educação, ciência e cultura (unesco). 2010. Citado na página 16.
- MARTINS, R. L. C.; LINHARES, M. P.; REIS, E. M. Mapas conceituais como instrumento de avaliação e aprendizagem de conceitos físicos sobre mecânica do voo. *Revista Brasileira de pesquisa em Educação em Ciências*, v. 9, n. 1, 2011. Citado 2 vezes nas páginas 36 e 37.

- MATURANA, H. R. *et al.* *Cognição, ciência e vida cotidiana*. [S.l.]: Ed. UFMG Belo Horizonte, 2001. Citado na página 36.
- MITRE, S. M. *et al.* Metodologias ativas de ensino-aprendizagem na formação profissional em saúde: debates atuais. *Ciência & saúde coletiva*, SciELO Public Health, v. 13, p. 2133–2144, 2008. Citado na página 18.
- MORAES, J. U. P.; SANTANA, R.; VIANA-BARBOSA, C. J. Avaliação baseada na aprendizagem significativa por meio de mapas conceituais. *Atas do VIII ENPEC, Campinas*, 2011. Citado 3 vezes nas páginas 36, 38 e 39.
- MOREIRA, M. A. Aprendizagem significativa: da visão clássica à visão crítica (meaningful learning: from the classical to the critical view). In: *Conferência de encerramento do V Encontro Internacional sobre Aprendizagem Significativa, Madrid, Espanha, setembro de* [S.l.: s.n.], 2006. Citado 2 vezes nas páginas 16 e 36.
- MOREIRA, M. A. Subsídios teóricos para o professor pesquisador em ensino de ciências: A teoria da aprendizagem significativa. *Porto Alegre-RS*, 2009. Citado 4 vezes nas páginas 23, 24, 25 e 26.
- MOREIRA, M. A. O mapa conceitual como instrumento de avaliação da aprendizagem. *Educação e Seleção*, n. 10, p. 17–34, 2013. Citado na página 21.
- MORESI, E. *et al.* Metodologia da pesquisa. *Brasília: Universidade Católica de Brasília*, v. 108, p. 24, 2003. Citado 2 vezes nas páginas 20 e 23.
- MÔRO, D. K. *et al.* Reconhecimento de entidades nomeadas em documentos de língua portuguesa. Araranguá, SC, 2018. Citado 2 vezes nas páginas 46 e 49.
- MULLER, E.; GRANATYR, J.; LESSING, O. R. Comparativo entre o algoritmo de luhn e o algoritmo gistsumm para sumarização de documentos. *Revista de Informática Teórica e Aplicada*, v. 22, n. 1, p. 75–94, 2015. Citado na página 57.
- NOTHMAN, J. *et al.* Learning multilingual named entity recognition from wikipedia. *Artificial Intelligence*, Elsevier, v. 194, p. 151–175, 2013. Citado na página 35.
- NOVAK, J. D. Uma teoria de educação. *São Paulo: Pioneira*, p. 55–73, 1981. Citado na página 36.
- NOVAK, J. D.; CAÑAS, A. J. The origins of the concept mapping tool and the continuing evolution of the tool. *Information visualization*, SAGE Publications Sage UK: London, England, v. 5, n. 3, p. 175–184, 2006. Citado 2 vezes nas páginas 27 e 30.
- NOVAK, J. D.; CAÑAS, A. J. The theory underlying concept maps and how to construct them. *Florida Institute for Human and Machine Cognition*, Citeseer, v. 1, p. 2006–2001, 2006. Citado na página 29.
- NOVAK, J. D.; CAÑAS, A. J. A teoria subjacente aos mapas conceituais e como elaborá-los e usá-los. *Práxis Educativa*, Universidade Estadual de Ponta Grossa, v. 5, n. 1, p. 9–29, 2010. Citado 2 vezes nas páginas 19 e 21.
- NOVAK, J. D.; GOWIN, D. B. *Learning How to Learn*. [S.l.]: Cambridge University Press, 1984. Citado 2 vezes nas páginas 31 e 42.

- NUNES, V. W. do N.; BESSA, R. C. Metodologias ativas apoiadas por recursos digitais: usando os aplicativos prezi e plickers. *Challenges 2017*, p. 25, 2018. Citado na página 18.
- PEREIRA, M. J. S. Processamento de linguagem natural para produtos de seguros. 2014. Citado na página 33.
- PEREIRA, S. d. L. *Processamento de linguagem natural*. [S.l.]: São Paulo: Universidade de São Paulo, 2011. Citado na página 32.
- PÉREZ, C. C. C.; VIEIRA, R. Mapas conceituais: geração e avaliação. In: *Anais do III Workshop em Tecnologia da Informação e da Linguagem Humana (TIL'2005)*. [S.l.: s.n.], 2005. p. 2158–2167. Citado 2 vezes nas páginas 39 e 45.
- PERIN, W. de A.; CURY, D.; MENEZES, C. S. de. Construindo mapas conceituais utilizando a abordagem imap. In: *XVII Congresso Internacional de Informática Educativa, TISE*. [S.l.: s.n.], 2012. Citado na página 53.
- PUENTES, R. V.; AQUINO, O. F. A aula universitária: resultados de um estudo empírico sobre o gerenciamento do tempo. *Linhas Críticas*, Universidade de Brasília, v. 14, n. 26, p. 111–129, 2008. Citado 2 vezes nas páginas 17 e 19.
- RIZKALLAH, J. The big (unstructured) data problem. *Forbes*. Retrieved on September, v. 5, p. 2017, 2017. Citado na página 31.
- RUIZ-MORENO, L. *et al.* Mapa conceitual: Ensaiando critérios de análise concept map: testing analytical criteria. *Ciência & Educação*, SciELO Brasil, v. 13, n. 3, p. 453–63, 2007. Citado 3 vezes nas páginas 30, 39 e 40.
- SADOSKI, M.; PAIVIO, A. *Imagery and text: A dual coding theory of reading and writing*. [S.l.]: Routledge, 2013. Citado na página 29.
- SETTE, B. S.; MARTINS, C. A. Pré-processamento textual para a extração de informação em bases de patentes. *Anais da Escola Regional de Informática da Sociedade Brasileira de Computação (SBC)–Regional de Mato Grosso*, v. 7, 2016. Citado 2 vezes nas páginas 46 e 48.
- TAPSCOTT, D.; WILLIAMS, A. D. Innovating the 21st-century university: It's time. *Educause review*, v. 45, n. 1, p. 16–29, 2010. Citado na página 17.
- TOFILON, P. J.; FIKE, J. R. The radioresponse of the central nervous system: a dynamic process. *Radiation research*, v. 153, n. 4, p. 357–370, 2000. Citado na página 30.
- VALENTE, J. A. Blended learning e as mudanças no ensino superior: a proposta da sala de aula invertida. *Educar em Revista*, Universidade Federal do Paraná, n. 4, p. 79–97, 2014. Citado na página 16.
- VALERIO, A.; LEAKE, D. B.; CAÑAS, A. J. Automatic classification of concept maps based on a topological taxonomy and its application to studying features of human-built maps. Tallinn University, 2008. Citado 3 vezes nas páginas 39, 43 e 44.
- VEKIRI, I. What is the value of graphical displays in learning? *Educational psychology review*, Springer, v. 14, n. 3, p. 261–312, 2002. Citado na página 29.

- VERGNAUD, G. La teoría de los campos conceptuales. *Recherches en didactique des mathématiques*, v. 10, n. 2, p. 3, 1990. Citado na página 36.
- VYGOTSKY, L. S. A formação social da mente. brasileira. *São Paulo, Martins*, 1988. Citado na página 23.
- VYGOTSKY, L. S. *The collected works of LS Vygotsky: Problems of the theory and history of psychology*. [S.l.]: Springer Science & Business Media, 1997. v. 3. Citado na página 37.
- YOUNG, T. *et al.* Recent trends in deep learning based natural language processing. *IEEE Computational Intelligence Magazine*, IEEE, v. 13, n. 3, p. 55–75, 2018. Citado na página 31.

Apêndices

APÊNDICE A – INSTRUÇÕES



Governo do Estado do Rio Grande do Norte
UNIVERSIDADE DO ESTADO DO RIO GRANDE DO NORTE – UERN
FACULDADE DE CIÊNCIAS EXATAS E NATURAIS – FANAT
PROGRAMA DE PÓS GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO- PPgCC
Campus Universitário-Br -110, KM-46-Costa e Silva- Mossoró/RN-Brasil
Fone/Fax: 84 3315-2196 – Ramal 2196 – e-mail: ppgcc@uern.br



Instruções:

O aluno deverá realizar a construção do Mapa Conceitual no tema sugerido “Processamento de Linguagem Natural – PLN” utilizando o software CmapTools. Esse trabalho deverá ser individual e baseado no material didático disponibilizado.

Base de Dados:

- Instalação do CmapTools: <<https://cmap.ihmc.us/cmaptools/cmaptools-download/>>
- Análise do material didático: <[Link](#)>
- Construção do Mapa Conceitual baseado no tutorial: <[Link](#)>

Ao concluir a tarefa o aluno deverá salvar o Mapa Conceitual no formato *Outline Cmap* (conforme tutorial) no padrão: Mapa_Nome_Aluno.

O Arquivo deverá ser enviado a bezerraneto@uern.br

APÊNDICE B – CLASSIFICAÇÕES GRAMATICAIS

```
[('pln', 'DET'), ('complicações', 'NOUN'), ('análise', 'NOUN'), ('fonética', 'ADJ'), ('\n', 'SPACE'), ('pln', 'DET'), ('complicações', 'NOUN'), ('expressões', 'ADJ'), ('do', 'ADP'), ('idioma', 'NOUN'), ('\n', 'SPACE'), ('pln', 'DET'), ('complicações', 'NOUN'), ('linguagem', 'NOUN'), ('não', 'ADV'), ('padrão', 'VERB'), ('\n', 'SPACE'), ('pln', 'ADV'), ('estuda', 'VERB'), ('capacidade', 'NOUN'), ('e', 'CCONJ'), ('limitação', 'NOUN'), ('de', 'ADP'), ('uma', 'DET'), ('máquina', 'NOUN'), ('\n', 'SPACE'), ('pln', 'ADV'), ('iniciado', 'VERB'), ('por', 'ADP'), ('alan', 'NOUN'), ('turing', 'NOUN'), ('\n', 'SPACE'), ('pln', 'ADV'), ('objetivo', 'ADJ'), ('normalizar', 'VERB'), ('\n', 'SPACE'), ('pln', 'ADV'), ('subárea', 'NOUN'), ('inteligência', 'ADJ'), ('artificial', 'ADJ'), ('\n', 'SPACE'), ('expressões', 'NOUN'), ('do', 'ADP'), ('idioma', 'NOUN'), ('tem-se', 'NOUN'), ('análise', 'NOUN'), ('morfológica', 'ADJ'), ('\n', 'SPACE'), ('expressões', 'NOUN'), ('do', 'ADP'), ('idioma', 'NOUN'), ('tem-se', 'ADV'), ('análise', 'NOUN'), ('semântica', 'ADJ'), ('\n', 'SPACE'), ('expressões', 'NOUN'), ('do', 'ADP'), ('idioma', 'NOUN'), ('tem-se', 'NOUN'), ('análise', 'NOUN'), ('sintática', 'ADJ'), ('\n', 'SPACE'), ('capacidade', 'NOUN'), ('e', 'CCONJ'), ('limitação', 'NOUN'), ('de', 'ADP'), ('uma', 'DET'), ('máquina', 'NOUN'), ('onde', 'ADV'), ('há', 'VERB'), ('dificuldade', 'NOUN'), ('de', 'ADP'), ('compreensão', 'NOUN'), ('da', 'ADP'), ('linguagem', 'NOUN'), ('escrita', 'VERB'), ('e', 'CCONJ'), ('falada', 'VERB'), ('\n', 'SPACE'), ('normalizar', 'VERB'), ('incluindo', 'VERB'), ('correção', 'NOUN'), ('ortográfica', 'ADJ'), ('\n', 'SPACE'), ('normalizar', 'VERB'), ('incluindo', 'VERB'), ('extração', 'NOUN'), ('de', 'ADP'), ('termos', 'NOUN'), ('\n', 'SPACE'), ('normalizar', 'VERB'), ('incluindo', 'VERB'), ('remoção', 'NOUN'), ('de', 'ADP'), ('numerais', 'NOUN'), ('\n', 'SPACE'), ('normalizar', 'VERB'), ('incluindo', 'VERB'), ('remoção', 'NOUN'), ('de', 'ADP'), ('stopwords', 'NOUN'), ('\n', 'SPACE'), ('normalizar', 'VERB'), ('incluindo', 'ADJ'), ('tokenização', 'NOUN'), ('\n', 'SPACE'), ('extração', 'NOUN'), ('de', 'ADP'), ('termos', 'NOUN'), ('abordagens', 'NOUN'), ('estatística', 'ADJ'), ('\n', 'SPACE'), ('extração', 'NOUN'), ('de', 'ADP'), ('termos', 'NOUN'), ('abordagens', 'NOUN'), ('híbrida', 'VERB'), ('\n', 'SPACE'), ('extração', 'NOUN'), ('de', 'ADP'), ('termos', 'NOUN'), ('abordagens', 'NOUN'), ('linguística', 'ADJ'), ('\n', 'SPACE'), ('inteligência', 'NOUN'), ('artificial', 'ADJ'), ('baseado', 'VERB'), ('na', 'ADP'), ('linguística', 'VERB'), ('\n', 'SPACE'), ('linguística', 'ADJ'), ('feito', 'VERB'), ('para', 'ADP'), ('processamento', 'NOUN'), ('em', 'ADP'), ('linguagem', 'NOUN'), ('natural', 'ADJ'), ('\n', 'SPACE'), ('processamento', 'NOUN'), ('em', 'ADP'), ('linguagem', 'NOUN'), ('natural', 'ADJ'), ('ferramentas', 'NOUN'), ('de', 'ADP'), ('auxílio', 'NOUN'), ('apache', 'ADJ'), ('lucene', 'ADJ'), ('\n', 'SPACE'), ('processamento', 'NOUN'), ('em', 'ADP'), ('linguagem', 'NOUN'), ('natural', 'ADJ'), ('ferramentas', 'NOUN'), ('de', 'ADP'), ('auxílio', 'NOUN'), ('em', 'ADP'), ('linguagem', 'NOUN'), ('natural', 'ADJ'), ('ferramentas', 'NOUN'), ('de', 'ADP'), ('auxílio', 'NOUN'), ('framenet', 'X'), ('processamento', 'NOUN'), ('em', 'ADP'), ('linguagem', 'NOUN'), ('natural', 'ADJ'), ('ferramentas', 'NOUN'), ('de', 'ADP'), ('auxílio', 'NOUN'), ('cogcomp', 'ADJ'), ('\n', 'SPACE'), ('processamento', 'NOUN'), ('em', 'ADP'), ('linguagem', 'NOUN'), ('natural', 'ADJ'), ('ferramentas', 'NOUN'), ('de', 'ADP'), ('auxílio', 'NOUN'), ('cogroo', 'ADJ'), ('\n', 'SPACE'), ('processamento', 'NOUN'), ('em', 'ADP'), ('linguagem', 'NOUN'), ('natural', 'ADJ'), ('ferramentas', 'NOUN'), ('de', 'ADP'), ('auxílio', 'NOUN'), ('framenet', 'X'), ('processamento', 'NOUN'), ('em', 'ADP'), ('linguagem', 'NOUN'), ('natural', 'ADJ'), ('ferramentas', 'NOUN'), ('de', 'ADP'), ('auxílio', 'NOUN'), ('stanford', 'VERB'), ('corenlp', 'ADV'), ('\n', 'SPACE'), ('processamento', 'NOUN'), ('em', 'ADP'), ('linguagem', 'NOUN'), ('natural', 'ADJ'), ('ferramentas', 'NOUN'), ('de', 'ADP'), ('auxílio', 'NOUN'), ('ultima', 'ADJ')]
```

Aluno Beta

```
[('pln', 'DET'), ('busca', 'ADJ'), ('soluções', 'NOUN'), ('\n', 'SPACE'), ('pln', 'ADV'), ('possibilita', 'VERB'), ('a', 'DET'), ('compreensão', 'NOUN'), ('de', 'ADP'), ('textos', 'NOUN'), ('\n', 'SPACE'), ('pln', 'ADV'), ('possui', 'VERB'), ('abordagem', 'NOUN'), ('estatística', 'ADJ'), ('\n', 'SPACE'), ('pln', 'ADV'), ('possui', 'VERB'), ('abordagem', 'NOUN'), ('híbrida', 'VERB'), ('\n', 'SPACE'), ('pln', 'ADV'), ('possui', 'VERB'), ('abordagem', 'NOUN'), ('linguística', 'ADJ'), ('\n', 'SPACE'), ('pln', 'ADV'), ('possui', 'VERB'), ('dificuldades', 'NOUN'), ('em', 'ADP'), ('conhecimento', 'NOUN'), ('\n', 'SPACE'), ('pln', 'ADV'), ('possui', 'VERB'), ('dificuldades', 'NOUN'), ('em', 'ADP'), ('expressões', 'NOUN'), ('idiomáticas', 'ADJ'), ('\n', 'SPACE'), ('pln', 'ADV'), ('possui', 'VERB'), ('dificuldades', 'NOUN'), ('em', 'ADP'), ('linguagem', 'NOUN'), ('padrão', 'ADJ'), ('\n', 'SPACE'), ('soluções', 'NOUN'), ('para', 'ADP'), ('problemas', 'NOUN'), ('computacionais', 'ADJ'), ('\n', 'SPACE'), ('a', 'DET'), ('compreensão', 'NOUN'), ('de', 'ADP'), ('textos', 'NOUN'), ('com', 'ADP'), ('aprendizagem', 'NOUN'), ('de', 'ADP'), ('conceitos', 'NOUN'), ('\n', 'SPACE'), ('a', 'DET'), ('compreensão', 'NOUN'), ('de', 'ADP'), ('textos', 'NOUN'), ('com', 'ADP'), ('interpretação', 'NOUN'), ('do', 'ADP'), ('sentido', 'NOUN'), ('\n', 'SPACE'), ('estatística', 'ADJ'), ('busca', 'NOUN'), ('unigramas', 'ADJ'), ('bigramas', 'NOUN'), ('ou', 'PUNCT'), ('trigramas', 'NOUN'), ('\n', 'SPACE'), ('estatística', 'ADJ'), ('realiza', 'ADJ'), ('análise', 'NOUN'), ('de', 'ADP'), ('frequência', 'NOUN'), ('de', 'ADP'), ('ocorrência', 'NOUN'), ('\n', 'SPACE'), ('análise', 'NOUN'), ('de', 'ADP'), ('frequência', 'NOUN'), ('de', 'ADP'), ('ocorrência', 'NOUN'), ('para', 'ADP'), ('extrair', 'VERB'), ('termos', 'NOUN'), ('\n', 'SPACE'), ('híbrida', 'VERB'), ('combina', 'NOUN'), ('as', 'DET'), ('abordagens', 'NOUN'), ('linguística', 'ADJ'), ('e', 'CCONJ'), ('estatística', 'ADJ'), ('\n', 'SPACE'), ('linguística', 'ADJ'), ('determina', 'VERB'), ('termos', 'NOUN'), ('candidatos', 'NOUN'), ('em', 'ADP'), ('linguística', 'ADJ'), ('identifica', 'ADJ'), ('sintagmas', 'NOUN'), ('nominais', 'ADJ'), ('\n', 'SPACE'), ('linguística', 'ADJ'), ('utiliza', 'VERB'), ('parsers', 'NOUN'), ('\n', 'SPACE'), ('termos', 'NOUN'), ('candidatos', 'NOUN'), ('para', 'ADP'), ('a', 'DET'), ('criação', 'NOUN'), ('de', 'ADP'), ('ontologias', 'NOUN'), ('\n', 'SPACE'), ('sintagmas', 'NOUN'), ('nominais', 'ADJ'), ('para', 'ADP'), ('extrair', 'VERB'), ('unidades', 'NOUN'), ('de', 'ADP'), ('sentido', 'NOUN'), ('\n', 'SPACE'), ('parsers', 'NOUN'), ('para', 'ADP'), ('fazer', 'VERB'), ('anotações', 'NOUN'), ('\n', 'SPACE'), ('fazer', 'VERB'), ('anotações', 'NOUN'), ('sendo', 'VERB'), ('elas', 'DET'), ('semânticas', 'ADJ'), ('\n', 'SPACE'), ('fazer', 'VERB'), ('anotações', 'NOUN'), ('sendo', 'VERB'), ('elas', 'VERB'), ('sintáticas', 'NOUN')]
```

Aluno Gama

```
[('pln', 'ADV'), ('possui', 'VERB'), ('abordagens', 'NOUN'), ('\n', 'SPACE'), ('pln', 'ADJ'), ('requer', 'ADJ'), ('tratamento', 'NOUN'), ('digital', 'ADJ'), ('\n', 'SPACE'), ('abordagens', 'NOUN'), ('podem', 'AUX'), ('ser', 'VERB'), ('estatística', 'ADJ'), ('\n', 'SPACE'), ('abordagens', 'NOUN'), ('podem', 'AUX'), ('ser', 'AUX'), ('hibrida', 'VERB'), ('\n', 'SPACE'), ('abordagens', 'NOUN'), ('podem', 'AUX'), ('ser', 'VERB'), ('linguística', 'ADJ'), ('\n', 'SPACE'), ('tratamento', 'NOUN'), ('digital', 'ADJ'), ('tem', 'VERB'), ('pré-processamentos', 'NOUN'), ('\n', 'SPACE'), ('pré-processamentos', 'ADJ'), ('podem', 'AUX'), ('ser', 'VERB'), ('normalização', 'NOUN'), ('\n', 'SPACE'), ('pré-processamentos', 'NOUN'), ('podem', 'AUX'), ('ser', 'VERB'), ('numerais', 'ADJ'), ('\n', 'SPACE'), ('pré-processamentos', 'ADJ'), ('podem', 'AUX'), ('ser', 'VERB'), ('ortografia', 'NOUN'), ('\n', 'SPACE'), ('pré-processamentos', 'NOUN'), ('podem', 'AUX'), ('ser', 'VERB'), ('stematização', 'NOUN'), ('e', 'CCONJ'), ('lematização', 'NOUN'), ('\n', 'SPACE'), ('pré-processamentos', 'NOUN'), ('podem', 'AUX'), ('ser', 'AUX'), ('stopwords', 'VERB'), ('\n', 'SPACE'), ('numerais', 'ADJ'), ('requer', 'ADJ'), ('normalização', 'NOUN'), ('\n', 'SPACE'), ('ortografia', 'VERB'), ('requer', 'ADJ'), ('normalização', 'NOUN'), ('\n', 'SPACE'), ('stematização', 'NOUN'), ('e', 'CCONJ'), ('lematização', 'NOUN'), ('requer', 'ADJ'), ('normalização', 'NOUN'), ('\n', 'SPACE'), ('stopwords', 'ADJ'), ('requer', 'VERB'), ('normalização', 'NOUN')]
```

Aluno Delta

```
[('o', 'DET'), ('pln', 'ADJ'), ('busca', 'ADJ'), ('soluções', 'NOUN'), ('para', 'ADP'), ('questões', 'NOUN'), ('computacionais', 'ADJ'), ('\n', 'SPACE'), ('o', 'DET'), ('pln', 'NOUN'), ('dedica-se', 'PROPN'), ('a', 'ADP'), ('propor', 'VERB'), ('e', 'CCONJ'), ('desenvolver', 'VERB'), ('sistemas', 'NOUN'), ('\n', 'SPACE'), ('para', 'ADP'), ('questões', 'NOUN'), ('computacionais', 'ADJ'), ('que', 'PRON'), ('requerem', 'VERB'), ('tratamento', 'NOUN'), ('computacional', 'ADJ'), ('\n', 'SPACE'), ('tratamento', 'NOUN'), ('computacional', 'ADJ'), ('que', 'PRON'), ('possui', 'VERB'), ('a', 'DET'), ('língua', 'NOUN'), ('natural', 'ADJ'), ('como', 'ADP'), ('objetivo', 'NOUN'), ('primario', 'ADJ'), ('\n', 'SPACE'), ('a', 'DET'), ('língua', 'NOUN'), ('natural', 'ADJ'), ('como', 'ADP'), ('objetivo', 'NOUN'), ('primario', 'ADJ'), ('seja', 'VERB'), ('ela', 'PRON'), ('escrita', 'VERB'), ('ou', 'CCONJ'), ('falada', 'VERB'), ('\n', 'SPACE'), ('a', 'ADP'), ('propor', 'VERB'), ('e', 'CCONJ'), ('desenvolver', 'VERB'), ('sistemas', 'NOUN'), ('que', 'PRON'), ('possuem', 'VERB'), ('a', 'DET'), ('língua', 'NOUN'), ('natural', 'ADJ'), ('como', 'ADP'), ('objetivo', 'NOUN'), ('primario', 'ADJ'), ('\n', 'SPACE'), ('a', 'DET'), ('língua', 'NOUN'), ('natural', 'ADJ'), ('como', 'ADP'), ('objetivo', 'NOUN'), ('primario', 'ADJ'), ('seja', 'VERB'), ('ela', 'PRON'), ('escrita', 'VERB'), ('ou', 'CCONJ'), ('falada', 'VERB')]
```

Aluno Épsilon

```
[('pln', 'ADJ'), ('busca', 'NOUN'), ('solucionar', 'VERB'), ('questões', 'NOUN'), ('computacionais', 'ADJ'), ('\n', 'SPACE'), ('pln', 'ADV'), ('fornece', 'VERB'), ('a', 'ADP'), ('os', 'DET'), ('computadores', 'SYM'), ('capacidade', 'ADJ'), ('entender', 'VERB'), ('textos', 'NOUN'), ('\n', 'SPACE'), ('pln', 'ADJ'), ('sub-área', 'NOUN'), ('da', 'ADP'), ('inteligência', 'NOUN'), ('artificial', 'ADJ'), ('\n', 'SPACE'), ('questões', 'NOUN'), ('computacionais', 'ADJ'), ('que', 'PRON'), ('inclui', 'VERB'), ('capacidade', 'NOUN'), ('entender', 'VERB'), ('textos', 'NOUN'), ('\n', 'SPACE'), ('capacidade', 'NOUN'), ('entender', 'VERB'), ('textos', 'NOUN'), ('criar', 'VERB'), ('resumos', 'NOUN'), ('\n', 'SPACE'), ('capacidade', 'NOUN'), ('entender', 'VERB'), ('textos', 'NOUN'), ('extrair', 'VERB'), ('informação', 'NOUN'), ('\n', 'SPACE'), ('capacidade', 'NOUN'), ('entender', 'VERB'), ('textos', 'NOUN'), ('interpretar', 'VERB'), ('sentidos', 'NOUN'), ('\n', 'SPACE'), ('capacidade', 'NOUN'), ('entender', 'VERB'), ('textos', 'NOUN'), ('necessidade', 'NOUN'), ('de', 'ADP'), ('pré-processamentos', 'NOUN'), ('\n', 'SPACE'), ('capacidade', 'NOUN'), ('entender', 'VERB'), ('textos', 'NOUN'), ('que', 'PRON'), ('significa', 'VERB'), ('analisar', 'VERB'), ('léxicamente', 'ADV'), ('\n', 'SPACE'), ('capacidade', 'NOUN'), ('entender', 'VERB'), ('textos', 'NOUN'), ('que', 'PRON'), ('significa', 'VERB'), ('analisar', 'VERB'), ('semânticamente', 'ADV'), ('\n', 'SPACE'), ('capacidade', 'NOUN'), ('entender', 'VERB'), ('textos', 'NOUN'), ('que', 'PRON'), ('significa', 'VERB'), ('analisar', 'VERB'), ('sentimentos', 'NOUN'), ('\n', 'SPACE'), ('capacidade', 'NOUN'), ('entender', 'VERB'), ('textos', 'NOUN'), ('que', 'PRON'), ('significa', 'VERB'), ('analisar', 'VERB'), ('sintaticamente', 'ADV'), ('\n', 'SPACE'), ('pré-processamentos', 'NOUN'), ('deixando', 'VERB'), ('apenas', 'ADV'), ('informações', 'NOUN'), ('relevantes', 'ADJ'), ('\n', 'SPACE'), ('capacidade', 'NOUN'), ('entender', 'VERB'), ('textos', 'NOUN'), ('criar', 'VERB'), ('resumos', 'NOUN'), ('\n', 'SPACE'), ('capacidade', 'NOUN'), ('entender', 'VERB'), ('textos', 'NOUN'), ('extrair', 'VERB'), ('informação', 'NOUN'), ('\n', 'SPACE'), ('capacidade', 'NOUN'), ('entender', 'VERB'), ('textos', 'NOUN'), ('interpretar', 'VERB'), ('sentidos', 'NOUN'), ('\n', 'SPACE'), ('capacidade', 'NOUN'), ('entender', 'VERB'), ('textos', 'NOUN'), ('necessidade', 'NOUN'), ('de', 'ADP'), ('pré-processamentos', 'NOUN'), ('\n', 'SPACE'), ('capacidade', 'NOUN'), ('entender', 'VERB'), ('textos', 'NOUN'), ('que', 'PRON'), ('significa', 'VERB'), ('analisar', 'VERB'), ('léxicamente', 'ADV'), ('\n', 'SPACE'), ('capacidade', 'NOUN'), ('entender', 'VERB'), ('textos', 'NOUN'), ('que', 'PRON'), ('significa', 'VERB'), ('analisar', 'VERB'), ('semânticamente', 'ADV'), ('\n', 'SPACE'), ('capacidade', 'NOUN'), ('entender', 'VERB'), ('textos', 'NOUN'), ('que', 'PRON'), ('significa', 'VERB'), ('analisar', 'VERB'), ('sentimentos', 'NOUN'), ('\n', 'SPACE'), ('capacidade', 'NOUN'), ('entender', 'VERB'), ('textos', 'NOUN'), ('que', 'PRON'), ('significa', 'VERB'), ('analisar', 'VERB'), ('sintaticamente', 'ADV'), ('\n', 'SPACE'), ('pré-processamentos', 'NOUN'), ('deixando', 'VERB'), ('apenas', 'ADV'), ('informações', 'NOUN'), ('relevantes', 'ADJ')]
```

Aluno Zeta

```
[('processamento', 'NOUN'), ('de', 'ADP'), ('linguagem', 'NOUN'), ('natural', 'ADJ'), ('modela', 'VERB'), ('a', 'DET'), ('linguagem', 'NOUN'), ('humana', 'ADJ'), ('\n', 'SPACE'), ('processamento', 'NOUN'), ('de', 'ADP'), ('linguagem', 'NOUN'), ('natural', 'ADJ'), ('possibilitou', 'VERB'), ('desenvolvimento', 'NOUN'), ('de', 'ADP'), ('aplicações', 'NOUN'), ('\n', 'SPACE'), ('processamento', 'NOUN'), ('de', 'ADP'), ('linguagem', 'NOUN'), ('natural', 'ADJ'), ('é', 'VERB'), ('uma', 'DET'), ('ferramenta', 'NOUN'), ('para', 'ADP'), ('ADP'), ('de', 'ADP'), ('linguagem', 'NOUN'), ('de', 'ADP'), ('linguagem', 'NOUN'), ('natural', 'ADJ'), ('é', 'VERB'), ('uma', 'DET'), ('subárea', 'NOUN'), ('da', 'ADP'), ('inteligência', 'NOUN'), ('artificial', 'ADJ'), ('\n', 'SPACE'), ('a', 'DET'), ('linguagem', 'NOUN'), ('humana', 'ADJ'), ('através', 'ADP'), ('máquina', 'NOUN'), ('da', 'ADP'), ('máquina', 'NOUN'), ('\n', 'SPACE'), ('a', 'DET'), ('linguagem', 'NOUN'), ('humana', 'ADJ'), ('para', 'ADP'), ('que', 'SCONJ'), ('a', 'DET'), ('compressão', 'NOUN'), ('e', 'CCONJ'), ('comunicação', 'NOUN'), ('\n', 'SPACE'), ('máquina', 'VERB'), ('por', 'ADP'), ('um', 'DET'), ('pré-processamento', 'NOUN'), ('\n', 'SPACE'), ('pré-processamento', 'ADJ'), ('por', 'ADP'), ('meio', 'NOUN'), ('da', 'ADP'), ('correção', 'NOUN'), ('ortográfica', 'ADJ'), ('\n', 'SPACE'), ('pré-processamento', 'NOUN'), ('por', 'ADP'), ('meio', 'NOUN'), ('da', 'ADP'), ('normalização', 'NOUN'), ('\n', 'SPACE'), ('pré-processamento', 'ADJ'), ('por', 'ADP'), ('meio', 'NOUN'), ('da', 'ADP'), ('remoção', 'NOUN'), ('de', 'ADP'), ('numerais', 'NOUN'), ('\n', 'SPACE'), ('pré-processamento', 'ADJ'), ('por', 'ADP'), ('meio', 'NOUN'), ('da', 'ADP'), ('remoção', 'NOUN'), ('de', 'ADP'), ('stopwords', 'NOUN'), ('\n', 'SPACE'), ('pré-processamento', 'ADJ'), ('por', 'ADP'), ('meio', 'NOUN'), ('da', 'ADP'), ('stemização', 'NOUN'), ('e', 'CCONJ'), ('lematização', 'NOUN'), ('\n', 'SPACE'), ('a', 'DET'), ('compressão', 'NOUN'), ('e', 'CCONJ'), ('comunicação', 'NOUN'), ('entre', 'ADP'), ('usuário', 'NOUN'), ('e', 'CCONJ'), ('máquina', 'NOUN'), ('\n', 'SPACE'), ('usuário', 'NOUN'), ('e', 'CCONJ'), ('máquina', 'NOUN'), ('seja', 'VERB'), ('semelhante', 'ADJ'), ('a', 'DET'), ('humana', 'NOUN'), ('\n', 'SPACE'), ('desenvolvimento', 'NOUN'), ('de', 'ADP'), ('aplicações', 'NOUN'), ('para', 'ADP'), ('análise', 'NOUN'), ('fonética', 'ADJ'), ('\n', 'SPACE'), ('desenvolvimento', 'NOUN'), ('de', 'ADP'), ('aplicações', 'NOUN'), ('para', 'ADP'), ('análise', 'NOUN'), ('morfológica', 'ADJ'), ('\n', 'SPACE'), ('desenvolvimento', 'NOUN'), ('de', 'ADP'), ('aplicações', 'NOUN'), ('para', 'ADP'), ('análise', 'NOUN'), ('semântica', 'ADJ'), ('\n', 'SPACE'), ('desenvolvimento', 'NOUN'), ('de', 'ADP'), ('aplicações', 'NOUN'), ('para', 'ADP'), ('análise', 'NOUN'), ('sintética', 'ADJ'), ('\n', 'SPACE'), ('fonética', 'ADJ'), ('aprimorando', 'VERB'), ('a', 'DET'), ('compressão', 'NOUN'), ('e', 'CCONJ'), ('comunicação', 'NOUN'), ('\n', 'SPACE'), ('fonética', 'ADJ'), ('aprimorando', 'VERB'), ('a', 'DET'), ('compressão', 'NOUN'), ('e', 'CCONJ'), ('comunicação', 'NOUN'), ('\n', 'SPACE'), ('semântica', 'NOUN'), ('\n', 'SPACE'), ('fonética', 'ADJ'), ('aprimorando', 'VERB'), ('sintética', 'NOUN'), ('\n', 'SPACE'), ('a', 'DET'), ('compressão', 'NOUN'), ('e', 'CCONJ'), ('comunicação', 'NOUN'), ('entre', 'ADP'), ('usuário', 'NOUN'), ('e', 'CCONJ'), ('máquina', 'NOUN'), ('\n', 'SPACE'), ('usuário', 'NOUN'), ('e', 'CCONJ'), ('máquina', 'NOUN'), ('seja', 'VERB'), ('semelhante', 'ADJ'), ('a', 'DET'), ('humana', 'NOUN'), ('para', 'ADP'), ('ferramenta', 'ADJ'), ('para', 'ADP'), ('melhor', 'ADV'), ('interpretar', 'VERB'), ('a', 'DET'), ('linguagem', 'NOUN'), ('humana', 'ADJ'), ('\n', 'SPACE'), ('a', 'DET'), ('linguagem', 'NOUN'), ('humana', 'ADJ'), ('através', 'NOUN'), ('da', 'ADP'), ('máquina', 'NOUN'), ('\n', 'SPACE'), ('a', 'DET'), ('linguagem', 'NOUN'), ('humana', 'ADJ'), ('para', 'ADP'), ('que', 'SCONJ'), ('a', 'DET'), ('compressão', 'NOUN'), ('e', 'CCONJ'), ('comunicação', 'NOUN'), ('\n', 'SPACE'), ('máquina', 'VERB'), ('por', 'ADP'), ('um', 'DET'), ('pré-processamento', 'NOUN'), ('\n', 'SPACE'), ('pré-processamento', 'ADJ'), ('por', 'ADP'), ('meio', 'NOUN'), ('da', 'ADP'), ('correção', 'NOUN'), ('ortográfica', 'ADJ'), ('\n', 'SPACE'), ('pré-processamento', 'NOUN'), ('por', 'ADP'), ('meio', 'NOUN'), ('da', 'ADP'), ('normalização', 'NOUN'), ('\n', 'SPACE'), ('pré-processamento', 'ADJ'), ('por', 'ADP'), ('meio', 'NOUN'), ('da', 'ADP'), ('remoção', 'NOUN'), ('de', 'ADP'), ('numerais', 'NOUN'), ('\n', 'SPACE'), ('pré-processamento', 'ADJ'), ('por', 'ADP'), ('meio', 'NOUN'), ('da', 'ADP'), ('remoção', 'NOUN'), ('de', 'ADP'), ('stopwords', 'NOUN'), ('\n', 'SPACE'), ('pré-processamento', 'ADJ'), ('por', 'ADP'), ('meio', 'NOUN'), ('da', 'ADP'), ('stemização', 'NOUN'), ('e', 'CCONJ'), ('lematização', 'NOUN'), ('\n', 'SPACE'), ('a', 'DET'), ('compressão', 'NOUN'), ('e', 'CCONJ'), ('comunicação', 'NOUN'), ('entre', 'ADP'), ('usuário', 'NOUN'), ('e', 'CCONJ'), ('máquina', 'NOUN'), ('\n', 'SPACE'), ('usuário', 'NOUN'), ('e', 'CCONJ'), ('máquina', 'NOUN'), ('seja', 'VERB'), ('semelhante', 'ADJ'), ('a', 'DET'), ('humana', 'NOUN')]]
```

Aluno Teta

```
[('pln', 'DET'), ('busca', 'ADJ'), ('soluções', 'NOUN'), ('computacionais', 'ADJ'), ('\n', 'SPACE'), ('pln', 'DET'), ('busca', 'NOUN'), ('compreender', 'VERB'), ('a', 'DET'), ('interpretação', 'NOUN'), ('\n', 'SPACE'), ('pln', 'DET'), ('busca', 'NOUN'), ('compreender', 'VERB'), ('linguagem', 'NOUN'), ('humana', 'ADJ'), ('\n', 'SPACE'), ('soluções', 'NOUN'), ('computacionais', 'ADJ'), ('que', 'PRON'), ('requerem', 'VERB'), ('tratamento', 'NOUN'), ('computacional', 'ADJ'), ('\n', 'SPACE'), ('tratamento', 'NOUN'), ('computacional', 'ADJ'), ('de', 'ADP'), ('uma', 'DET'), ('língua', 'NOUN'), ('natural', 'ADJ'), ('\n', 'SPACE'), ('língua', 'NOUN'), ('natural', 'ADJ'), ('que', 'PRON'), ('é', 'VERB'), ('escrita', 'VERB'), ('\n', 'SPACE'), ('língua', 'NOUN'), ('natural', 'ADJ'), ('que', 'PRON'), ('é', 'AUX'), ('falada', 'VERB'), ('\n', 'SPACE'), ('a', 'DET'), ('interpretação', 'NOUN'), ('relacionada', 'VERB'), ('a', 'DET'), ('linguagem', 'NOUN'), ('humana', 'ADJ'), ('\n', 'SPACE'), ('linguagem', 'NOUN'), ('humana', 'ADJ'), ('objetivando', 'VERB'), ('que', 'SCONJ'), ('computador', 'NOUN'), ('compreenda', 'VERB'), ('linguagem', 'NOUN'), ('\n', 'SPACE'), ('computador', 'NOUN'), ('compreenda', 'VERB'), ('linguagem', 'NOUN'), ('através', 'ADV'), ('de', 'ADP'), ('aplicações', 'NOUN'), ('\n', 'SPACE'), ('aplicações', 'NOUN'), ('com', 'ADP'), ('análise', 'NOUN'), ('fonética', 'ADJ'), ('\n', 'SPACE'), ('aplicações', 'NOUN'), ('com', 'ADP'), ('análise', 'NOUN'), ('morfológica', 'ADJ'), ('\n', 'SPACE'), ('aplicações', 'NOUN'), ('com', 'ADP'), ('análise', 'NOUN'), ('semântica', 'ADJ'), ('aplicações', 'NOUN'), ('com', 'ADP'), ('análise', 'NOUN'), ('sintática', 'ADJ'), ('\n', 'SPACE'), ('aplicações', 'NOUN'), ('que', 'PRON'), ('necessitam', 'VERB'), ('de', 'ADP'), ('tratamento', 'NOUN'), ('computacional', 'ADJ'), ('\n', 'SPACE'), ('tratamento', 'NOUN'), ('computacional', 'ADJ'), ('de', 'ADP'), ('uma', 'DET'), ('língua', 'NOUN'), ('natural', 'ADJ'), ('\n', 'SPACE'), ('língua', 'NOUN'), ('natural', 'ADJ'), ('que', 'PRON'), ('é', 'VERB'), ('escrita', 'VERB'), ('\n', 'SPACE'), ('língua', 'NOUN'), ('natural', 'ADJ'), ('que', 'PRON'), ('é', 'AUX'), ('falada', 'VERB')]]
```

Aluno Lambda

```
[('processamento', 'NOUN'), ('de', 'ADP'), ('linguagem', 'NOUN'), ('natural', 'ADJ'), ('busca', 'NOUN'), ('solução', 'NOUN'), ('para', 'ADP'), ('questão', 'NOUN'), ('computacional', 'ADJ'), ('\n', 'SPACE'), ('processamento', 'NOUN'), ('de', 'ADP'), ('linguagem', 'NOUN'), ('natural', 'ADJ'), ('desenvolve', 'VERB'), ('sistema', 'NOUN'), ('operacional', 'ADJ'), ('\n', 'SPACE'), ('processamento', 'NOUN'), ('de', 'ADP'), ('linguagem', 'NOUN'), ('natural', 'ADJ'), ('reduz', 'VERB'), ('vocabulário', 'ADJ'), ('\n', 'SPACE'), ('solução', 'NOUN'), ('para', 'ADP'), ('questão', 'NOUN'), ('computacional', 'ADJ'), ('utiliza', 'VERB'), ('linguagem', 'NOUN'), ('escrita', 'ADJ'), ('ou', 'CCONJ'), ('falada', 'VERB'), ('\n', 'SPACE'), ('sistema', 'NOUN'), ('operacional', 'ADJ'), ('com', 'ADP'), ('linguagem', 'NOUN'), ('natural', 'ADJ'), ('\n', 'SPACE'), ('linguagem', 'NOUN'), ('natural', 'ADJ'), ('como', 'ADP'), ('objeto', 'NOUN'), ('primário', 'ADJ'), ('\n', 'SPACE'), ('vocabulário', 'ADJ'), ('para', 'ADP'), ('linguagem', 'NOUN'), ('escrita', 'VERB'), ('ou', 'CCONJ'), ('falada', 'VERB')]]
```

Aluno Sigma

APÊNDICE C – AVALIAÇÕES DO ESPECIALISTA

IDENTIFICAÇÃO	α alfa			ANÁLISE FERRAMENTA			ANÁLISE ESPECIALISTA					
	NÍVEL	CONCEITO 1	VERBO DE LIGAÇÃO	CONCEITO 2	ESTRUTURA	GRAMÁTICA	CONTEXTO	ESTRUTURA	GRAMÁTICA	CONTEXTO		
N0	Processamento de linguagem natural	dedica-se	Desenvolver sistemas operacionais	SIM	NÃO		SIM	SIM				
N0	Processamento de linguagem natural	Sub-área	Inteligência Artificial	SIM	NÃO		SIM	NÃO				
N1	Desenvolver sistemas operacionais	Entender	Através do pre-processamento	SIM	SIM	NÃO	SIM	SIM	NÃO			
N1	Desenvolver sistemas operacionais	Entender	Seres humanos	SIM	SIM	NÃO	SIM	SIM	NÃO			
N1	Inteligência Artificial	Estuda	Capacidade da máquina	SIM	SIM	SIM	SIM	SIM	SIM			
N2	Através do pre-processamento	Verificando	Estrutura da língua	SIM	SIM	SIM	SIM	SIM	SIM			
N2	Capacidade da máquina	Entender	Através do pre-processamento	SIM	SIM	SIM	SIM	SIM	SIM			
N2	Capacidade da máquina	Entender	Seres humanos	SIM	SIM	SIM	SIM	SIM	SIM			
N3	Estrutura da língua	Reduzindo	Vocabulário	SIM	SIM	SIM	SIM	SIM	NÃO			
N3	Através do pre-processamento	Verificando	Estrutura da língua	SIM	SIM	SIM	SIM	SIM	SIM			
N4	Vocabulário	Utilizando	Correção ortográfica	SIM	SIM	SIM	SIM	SIM	SIM			
N4	Vocabulário	Utilizando	Normalização	SIM	SIM	SIM	SIM	SIM	SIM			
N4	Vocabulário	Utilizando	Remoção de numerais	SIM	SIM	SIM	SIM	SIM	SIM			
N4	Vocabulário	Utilizando	Remoção de stopwords	SIM	SIM	SIM	SIM	SIM	SIM			
N4	Vocabulário	Utilizando	Stemização e lematização	SIM	SIM	SIM	SIM	SIM	SIM			
N4	Estrutura da língua	Reduzindo	Vocabulário	SIM	SIM	SIM	SIM	SIM	NÃO			
N5	Remoção de stopwords	Remove	Palavras frequentes	SIM	NÃO		SIM	SIM				
N5	Stemização e lematização	Reduz	Palavra ao radical	SIM	SIM	SIM	SIM	SIM	SIM			
N5	Stemização e lematização	Reduz	Palavra ao seu lema	SIM	SIM	SIM	SIM	SIM	SIM			
N5	Vocabulário	Utilizando	Correção ortográfica	SIM	SIM	SIM	SIM	SIM	SIM			
N5	Vocabulário	Utilizando	Normalização	SIM	SIM	SIM	SIM	SIM	SIM			
N5	Vocabulário	Utilizando	Remoção de numerais	SIM	SIM	SIM	SIM	SIM	SIM			
N5	Vocabulário	Utilizando	Remoção de stopwords	SIM	SIM	SIM	SIM	SIM	SIM			
N5	Vocabulário	Utilizando	Stemização e lematização	SIM	SIM	SIM	SIM	SIM	SIM			
N6	Remoção de stopwords	Remove	Palavras frequentes	SIM	NÃO		SIM	SIM				
N6	Stemização e lematização	Reduz	Palavra ao radical	SIM	SIM	SIM	SIM	SIM	SIM			
N6	Stemização e lematização	Reduz	Palavra ao seu lema	SIM	SIM	SIM	SIM	SIM	SIM			
							27	24	21			
							100%	89%	91%			
			ESTRUTURA			GRAMÁTICA	CONTEXTO	ESTRUTURA			GRAMÁTICA	CONTEXTO
			ANÁLISE DA ARQUITETURA			ANÁLISE ESPECIALISTA						

Scoring Criteria for Concept Maps	96	85
	11	

Aluno Alfa

IDENTIFICAÇÃO	β beta			ANÁLISE FERRAMENTA			ANÁLISE ESPECIALISTA					
	NÍVEL	CONCEITO 1	VERBO DE LIGAÇÃO	CONCEITO 2	ESTRUTURA	GRAMÁTICA	CONTEXTO	ESTRUTURA	GRAMÁTICA	CONTEXTO		
N0	PLN	Comparações	Análise Fonética	SIM	NÃO		SIM	NÃO				
N0	PLN	Comparações	Expressões do idioma	SIM	NÃO		SIM	NÃO				
N0	PLN	Comparações	Linguagem não padrão	SIM	NÃO		SIM	NÃO				
N0	PLN	Estuda	Capacidade e limitação de uma Máquina	SIM	SIM	SIM	SIM	SIM	SIM			
N0	PLN	Iniciado por	Alan Turing	SIM	SIM	SIM	SIM	SIM	SIM			
N0	PLN	Objetivo	Normalizar	SIM	NÃO		SIM	NÃO				
N0	PLN	Subárea	Inteligência Artificial	SIM	NÃO		SIM	NÃO				
N1	Expressões do idioma	tem-se	Análise Morfológica	SIM	NÃO		SIM	SIM				
N1	Expressões do idioma	tem-se	Análise Semântica	SIM	NÃO		SIM	SIM				
N1	Expressões do idioma	tem-se	Análise Sintática	SIM	NÃO		SIM	SIM				
N1	Capacidade e limitação de uma Máquina	Onde há	Dificuldade de Compreensão da Linguagem escrita e falada	SIM	SIM	SIM	SIM	NÃO	SIM			
N1	Normalizar	Incluindo	Correção Ortográfica	SIM	SIM	NÃO	SIM	SIM	NÃO			
N1	Normalizar	Incluindo	Extração de Termos	SIM	SIM	SIM	SIM	SIM	SIM			
N1	Normalizar	Incluindo	Remoção de Numerais	SIM	SIM	NÃO	SIM	SIM	SIM			
N1	Normalizar	Incluindo	Remoção de Stopwords	SIM	SIM	NÃO	SIM	SIM	SIM			
N1	Normalizar	Incluindo	Tokenização	SIM	SIM	NÃO	SIM	SIM	SIM			
N1	Inteligência Artificial	Baseado na	Linguística	SIM	SIM	NÃO	SIM	SIM	NÃO			
N2	Extração de Termos	Abordagens	Estatística	SIM	NÃO		SIM	SIM				
N2	Extração de Termos	Abordagens	Híbrida	SIM	NÃO		SIM	SIM				
N2	Extração de Termos	Abordagens	Linguística	SIM	NÃO		SIM	SIM				
N2	Linguística	Feito Para	Processamento em Linguagem Natural	SIM	SIM	SIM	SIM	NÃO	SIM			
N3	Processamento em Linguagem Natural	Ferramentas de Auxílio	Apache Lucene	SIM	SIM		SIM	NÃO				
N3	Processamento em Linguagem Natural	Ferramentas de Auxílio	ARK	SIM	NÃO		SIM	NÃO				
N3	Processamento em Linguagem Natural	Ferramentas de Auxílio	COGCOMP	SIM	NÃO		SIM	NÃO				
N3	Processamento em Linguagem Natural	Ferramentas de Auxílio	CoGROO	SIM	NÃO		SIM	NÃO				
N3	Processamento em Linguagem Natural	Ferramentas de Auxílio	FrameNET	SIM	NÃO		SIM	NÃO				
N3	Processamento em Linguagem Natural	Ferramentas de Auxílio	Freeling	SIM	NÃO		SIM	NÃO				
N3	Processamento em Linguagem Natural	Ferramentas de Auxílio	Gate	SIM	NÃO		SIM	NÃO				
N3	Processamento em Linguagem Natural	Ferramentas de Auxílio	LX-Center	SIM	NÃO		SIM	NÃO				
N3	Processamento em Linguagem Natural	Ferramentas de Auxílio	Natural Language Toolkit	SIM	NÃO		SIM	NÃO				
N3	Processamento em Linguagem Natural	Ferramentas de Auxílio	openNLP	SIM	NÃO		SIM	NÃO				
N3	Processamento em Linguagem Natural	Ferramentas de Auxílio	Stanford CoreNLP	SIM	NÃO		SIM	NÃO				
N3	Processamento em Linguagem Natural	Ferramentas de Auxílio	UIMA	SIM	NÃO		SIM	NÃO				
							33	24	7			
							100%	73%	70%			
			ESTRUTURA			GRAMÁTICA	CONTEXTO	ESTRUTURA			GRAMÁTICA	CONTEXTO
			ANÁLISE FERRAMENTA			ANÁLISE ESPECIALISTA						

Scoring Criteria for Concept Maps	30	5
	25	

Aluno Beta

IDENTIFICAÇÃO		δ delta		ANÁLISE FERRAMENTA			ANÁLISE ESPECIALISTA			
NÍVEL	CONCEITO 1	VERBO DE LIGAÇÃO	CONCEITO 2	ESTRUTURA	GRAMÁTICA	CONTEXTO	ESTRUTURA	GRAMÁTICA	CONTEXTO	
N0	PLN	possui	abordagens	SIM	SIM	SIM	SIM	SIM	SIM	
N0	PLN	Requer	Tratamento digital	SIM	NÃO		SIM	SIM		
N1	abordagens	podem ser	estatística	SIM	SIM	SIM	SIM	SIM	SIM	
N1	abordagens	podem ser	híbrida	SIM	SIM	SIM	SIM	SIM	SIM	
N1	abordagens	podem ser	linguística	SIM	SIM	SIM	SIM	SIM	SIM	
N1	Tratamento digital	tem	pré-processamentos	SIM	SIM	SIM	SIM	SIM	NÃO	
N2	pré-processamentos	podem ser	normalização	SIM	SIM	SIM	SIM	SIM	NÃO	
N2	pré-processamentos	podem ser	numerais	SIM	SIM	SIM	SIM	SIM	NÃO	
N2	pré-processamentos	podem ser	ortografia	SIM	SIM	NÃO	SIM	SIM	NÃO	
N2	pré-processamentos	podem ser	Stematização e lematização	SIM	SIM	SIM	SIM	SIM	SIM	
N2	pré-processamentos	podem ser	stopwords	SIM	SIM	SIM	SIM	SIM	SIM	
N3	numerais	requer	normalização	SIM	NÃO		SIM	SIM		
N3	ortografia	requer	normalização	SIM	NÃO		SIM	SIM		
N3	Stematização e lematização	requer	normalização	SIM	NÃO		SIM	SIM		
N3	stopwords	requer	normalização	SIM	NÃO		SIM	SIM		
							15	10	7	
							AVALIAÇÃO	100%	67%	70%
				ESTRUTURA	GRAMÁTICA	CONTEXTO	ESTRUTURA	GRAMÁTICA	CONTEXTO	
				ANÁLISE FERRAMENTA			ANÁLISE ESPECIALISTA			

Scoring Criteria for Concept Maps	46	29
	17	

Aluno Delta

IDENTIFICAÇÃO		ε épsilon		ANÁLISE FERRAMENTA			ANÁLISE ESPECIALISTA			
NÍVEL	CONCEITO 1	VERBO DE LIGAÇÃO	CONCEITO 2	ESTRUTURA	GRAMÁTICA	CONTEXTO	ESTRUTURA	GRAMÁTICA	CONTEXTO	
N0	O PLN	busca soluções	para questões computacionais	SIM	NÃO		SIM	SIM		
N0	O PLN	dedica-se	a propor e desenvolver sistemas	SIM	NÃO		SIM	SIM		
N1	para questões computacionais	que requerem	tratamento computacional	SIM	SIM	SIM	SIM	SIM	SIM	
N1	a propor e desenvolver sistemas	que possuem	a língua natural como objetivo primario	SIM	SIM	NÃO	SIM	SIM	NÃO	
N2	tratamento computacional	que possui	a língua natural como objetivo primario	SIM	SIM	SIM	SIM	SIM	NÃO	
N2	a língua natural como objetivo primario	seja ela	escrita ou falada	SIM	SIM	SIM	SIM	SIM	SIM	
N3	a língua natural como objetivo primario	seja ela	escrita ou falada	SIM	SIM	SIM	SIM	SIM	SIM	
							7	5	4	
							AVALIAÇÃO	100%	71%	80%
				ESTRUTURA	GRAMÁTICA	CONTEXTO	ESTRUTURA	GRAMÁTICA	CONTEXTO	
				ANÁLISE FERRAMENTA			ANÁLISE ESPECIALISTA			

Scoring Criteria for Concept Maps	27	19
	8	

Aluno Épsilon

IDENTIFICAÇÃO		γ gama		ANÁLISE FERRAMENTA			ANÁLISE ESPECIALISTA			
NÍVEL	CONCEITO 1	VERBO DE LIGAÇÃO	CONCEITO 2	ESTRUTURA	GRAMÁTICA	CONTEXTO	ESTRUTURA	GRAMÁTICA	CONTEXTO	
N0	PLN	Busca	Soluções	SIM	NÃO		SIM	SIM		
N0	PLN	Possibilita	A compreensão de textos	SIM	SIM	SIM	SIM	SIM	SIM	
N0	PLN	Possui abordagem	Estatística	SIM	SIM	SIM	SIM	SIM	SIM	
N0	PLN	Possui abordagem	Híbrida	SIM	SIM	SIM	SIM	SIM	SIM	
N0	PLN	Possui abordagem	Linguística	SIM	SIM	SIM	SIM	SIM	SIM	
N0	PLN	Possui dificuldades em	Conhecimento	SIM	SIM	SIM	SIM	SIM	SIM	
N0	PLN	Possui dificuldades em	Expressões idiomáticas	SIM	SIM	SIM	SIM	SIM	SIM	
N0	PLN	Possui dificuldades em	Linguagem padrão	SIM	SIM	SIM	SIM	SIM	NÃO	
N1	Soluções	Para	Problemas computacionais	SIM	NÃO		SIM	NÃO		
N1	A compreensão de textos	Com	Aprendizagem de conceitos	SIM	NÃO		SIM	NÃO		
N1	A compreensão de textos	Com	Interpretação do sentido	SIM	NÃO		SIM	NÃO		
N1	Estatística	Busca	Unigramas, bigramas ou trigramas	SIM	NÃO		SIM	SIM		
N1	Estatística	Realiza	Análise de frequência de ocorrência	SIM	SIM	SIM	SIM	SIM	SIM	
N1	Híbrida	Combina	As abordagens (linguística e estatística)	SIM	SIM	SIM	SIM	SIM	SIM	
N1	Linguística	Determina	Termos candidatos	SIM	SIM	SIM	SIM	SIM	SIM	
N1	Linguística	Identifica	Sintagmas nominais	SIM	SIM	SIM	SIM	SIM	SIM	
N1	Linguística	Utiliza	Parsers	SIM	SIM	SIM	SIM	SIM	SIM	
N2	Análise de frequência de ocorrência	Para	Extrair termos	SIM	NÃO		SIM	NÃO		
N2	Termos candidatos	Para	A criação de ontologias	SIM	NÃO		SIM	NÃO		
N2	Sintagmas nominais	Para extrair	Unidades de sentido	SIM	SIM	SIM	SIM	SIM	SIM	
N2	Parsers	Para	Fazer anotações	SIM	NÃO		SIM	NÃO		
N3	Fazer anotações	Sendo elas	Semânticas	SIM	SIM	SIM	SIM	NÃO	NÃO	
N3	Fazer anotações	Sendo elas	Sintáticas	SIM	SIM	SIM	SIM	NÃO	NÃO	
							23	19	12	
							AVALIAÇÃO	100%	83%	80%
				ESTRUTURA	GRAMÁTICA	CONTEXTO	ESTRUTURA	GRAMÁTICA	CONTEXTO	
				ANÁLISE FERRAMENTA			ANÁLISE ESPECIALISTA			

Scoring Criteria for Concept Maps	65	65
	0	

Aluno Gama

IDENTIFICAÇÃO	λ lambda			ANÁLISE FERRAMENTA			ANÁLISE ESPECIALISTA		
	NÍVEL	CONCEITO 1	VERBO DE LIGAÇÃO	CONCEITO 2	ESTRUTURA	GRAMÁTICA	CONTEXTO	ESTRUTURA	GRAMÁTICA
N0	PLN	busca	Soluções computacionais	SIM	NÃO		SIM	SIM	
N0	PLN	busca compreender	A interpretação	SIM	SIM	SIM	SIM	SIM	SIM
N0	PLN	busca compreender	Linguagem humana	SIM	SIM	SIM	SIM	SIM	SIM
N1	Soluções computacionais	que requerem	Tratamento computacional	SIM	SIM	SIM	SIM	SIM	SIM
N1	A interpretação	relacionada a	Linguagem humana	SIM	SIM	SIM	SIM	SIM	SIM
N1	Linguagem humana	objetivando que	Computador compreenda linguagem	SIM	SIM	SIM	SIM	SIM	SIM
N2	Tratamento computacional	de uma	Lingua natural	SIM	NÃO		SIM	NÃO	
N2	Linguagem humana	objetivando que	Computador compreenda linguagem	SIM	SIM	SIM	SIM	SIM	SIM
N2	Computador compreenda linguagem	através de	Aplicações	SIM	NÃO		SIM	SIM	
N3	Lingua natural	que é	Escrita	SIM	SIM	SIM	SIM	SIM	SIM
N3	Lingua natural	que é	Falada	SIM	SIM	SIM	SIM	SIM	SIM
N3	Computador compreenda linguagem	através de	Aplicações	SIM	NÃO		SIM	SIM	
N3	Aplicações	com análise	Fonética	SIM	NÃO		SIM	SIM	
N3	Aplicações	com análise	Morfológica	SIM	NÃO		SIM	SIM	
N3	Aplicações	com análise	Semântica	SIM	NÃO		SIM	SIM	
N3	Aplicações	com análise	Sintática	SIM	NÃO		SIM	SIM	
N3	Aplicações	que necessitam de	Tratamento computacional	SIM	SIM	SIM	SIM	SIM	SIM
N4	Aplicações	com análise	Fonética	SIM	NÃO		SIM	SIM	
N4	Aplicações	com análise	Morfológica	SIM	NÃO		SIM	SIM	
N4	Aplicações	com análise	Semântica	SIM	NÃO		SIM	SIM	
N4	Aplicações	com análise	Sintática	SIM	NÃO		SIM	SIM	
N4	Aplicações	que necessitam de	Tratamento computacional	SIM	SIM	SIM	SIM	SIM	SIM
N4	Tratamento computacional	de uma	Lingua natural	SIM	NÃO		SIM	NÃO	
N5	Tratamento computacional	de uma	Lingua natural	SIM	NÃO		SIM	NÃO	
N5	Lingua natural	que é	Escrita	SIM	SIM	SIM	SIM	SIM	SIM
N5	Lingua natural	que é	Falada	SIM	SIM	SIM	SIM	SIM	SIM
N6	Lingua natural	que é	Escrita	SIM	SIM	SIM	SIM	SIM	SIM
N6	Lingua natural	que é	Falada	SIM	SIM	SIM	SIM	SIM	SIM
							28	17	14
							100%	61%	100%
ESTRUTURA GRAMÁTICA CONTEXTO							ESTRUTURA GRAMÁTICA CONTEXTO	ESTRUTURA GRAMÁTICA CONTEXTO	ESTRUTURA GRAMÁTICA CONTEXTO
ANÁLISE FERRAMENTA							ANÁLISE ESPECIALISTA		

Scoring Criteria for Concept Maps	77	77
	0	

Aluno Lambda

IDENTIFICAÇÃO	σ sigma			ANÁLISE FERRAMENTA			ANÁLISE ESPECIALISTA		
	NÍVEL	CONCEITO 1	VERBO DE LIGAÇÃO	CONCEITO 2	ESTRUTURA	GRAMÁTICA	CONTEXTO	ESTRUTURA	GRAMÁTICA
N0	Processamento de Linguagem Natural	Busca	Solução para questão computacional	SIM	NÃO		SIM	SIM	
N0	Processamento de Linguagem Natural	Desenvolve	Sistema operacional	SIM	SIM	NÃO	SIM	SIM	NÃO
N0	Processamento de Linguagem Natural	Reduz	Vocabulario	SIM	SIM	SIM	SIM	SIM	NÃO
N1	Solução para questão computacional	Utiliza	Linguagem Escrita ou falada	SIM	SIM	SIM	SIM	SIM	NÃO
N1	Sistema operacional	com	linguagem natural	SIM	NÃO		SIM	NÃO	
N1	Vocabulario	para	Linguagem Escrita ou falada	SIM	NÃO		SIM	NÃO	
N2	linguagem natural	Como	Objeto primario	SIM	NÃO		SIM	NÃO	
							7	6	1
							100%	86%	33%
ESTRUTURA GRAMÁTICA CONTEXTO							ESTRUTURA GRAMÁTICA CONTEXTO	ESTRUTURA GRAMÁTICA CONTEXTO	ESTRUTURA GRAMÁTICA CONTEXTO
ANÁLISE FERRAMENTA							ANÁLISE ESPECIALISTA		

Scoring Criteria for Concept Maps	11	8
	3	

Aluno Sigma

IDENTIFICAÇÃO	θ teta			ANÁLISE FERRAMENTA			ANÁLISE ESPECIALISTA		
	NÍVEL	CONCEITO 1	VERBO DE LIGAÇÃO	CONCEITO 2	ESTRUTURA	GRAMÁTICA	CONTEXTO	ESTRUTURA	GRAMÁTICA
N0	Processamento de Linguagem Natural	modela	a linguagem humana	SIM	NÃO		SIM	SIM	SIM
N0	Processamento de Linguagem Natural	possibilitou	desenvolvimento de aplicações	SIM	SIM	SIM	SIM	SIM	SIM
N0	Processamento de Linguagem Natural	é uma	ferramenta para melhor	SIM	SIM		SIM	SIM	NÃO
N0	Processamento de Linguagem Natural	é uma	subárea da Inteligência Artificial	SIM	SIM	SIM	SIM	SIM	SIM
N1	a linguagem humana	através da	máquina	SIM	NÃO		SIM	SIM	
N1	a linguagem humana	para que	a compreensão e comunicação	SIM	NÃO		SIM	NÃO	
N1	desenvolvimento de aplicações	para análise	fonética	SIM	NÃO		SIM	SIM	
N1	desenvolvimento de aplicações	para análise	morfológica	SIM	NÃO		SIM	SIM	
N1	desenvolvimento de aplicações	para análise	semântica	SIM	NÃO		SIM	SIM	
N1	desenvolvimento de aplicações	para análise	sintética	SIM	NÃO		SIM	SIM	
N1	ferramenta para melhor	interpretar	a linguagem humana	SIM	SIM	SIM	SIM	SIM	NÃO
N2	máquina	por um	pré-processamento	SIM	NÃO		SIM	NÃO	
N2	a compreensão e comunicação	entre	usuário e máquina	SIM	NÃO		SIM	NÃO	
N2	fonética	aprimorando	a compreensão e comunicação	SIM	SIM	NÃO	SIM	SIM	NÃO
N2	fonética	aprimorando	morfológica	SIM	SIM	NÃO	SIM	SIM	NÃO
N2	fonética	aprimorando	semântica	SIM	SIM	SIM	SIM	SIM	SIM
N2	fonética	aprimorando	sintética	SIM	SIM	NÃO	SIM	SIM	NÃO
N2	a linguagem humana	através da	máquina	SIM	NÃO		SIM	SIM	
N2	a linguagem humana	para que	a compreensão e comunicação	SIM	NÃO		SIM	NÃO	
N3	pré-processamento	por meio da	Correção Ortográfica	SIM	NÃO		SIM	NÃO	
N3	pré-processamento	por meio da	Normalização	SIM	NÃO		SIM	NÃO	
N3	pré-processamento	por meio da	Remoção de Numerais	SIM	NÃO		SIM	NÃO	
N3	pré-processamento	por meio da	Remoção de Stopwords	SIM	NÃO		SIM	NÃO	
N3	pré-processamento	por meio da	Stemização e Lematização	SIM	NÃO		SIM	NÃO	
N3	usuário e máquina	seja	semelhante a humana	SIM	SIM	SIM	SIM	SIM	SIM
N3	a compreensão e comunicação	entre	usuário e máquina	SIM	NÃO		SIM	NÃO	
N3	máquina	por um	pré-processamento	SIM	NÃO		SIM	NÃO	
N3	a compreensão e comunicação	entre	usuário e máquina	SIM	NÃO		SIM	NÃO	
N4	usuário e máquina	seja	semelhante a humana	SIM	SIM	SIM	SIM	SIM	SIM
N4	pré-processamento	por meio da	Correção Ortográfica	SIM	NÃO		SIM	NÃO	
N4	pré-processamento	por meio da	Normalização	SIM	NÃO		SIM	NÃO	
N4	pré-processamento	por meio da	Remoção de Numerais	SIM	NÃO		SIM	NÃO	
N4	pré-processamento	por meio da	Remoção de Stopwords	SIM	NÃO		SIM	NÃO	
N4	pré-processamento	por meio da	Stemização e Lematização	SIM	NÃO		SIM	NÃO	
N4	usuário e máquina	seja	semelhante a humana	SIM	SIM	SIM	SIM	SIM	SIM
							35	28	9
							100%	80%	82%
			ANÁLISE FERRAMENTA			ANÁLISE ESPECIALISTA			

Scoring Criteria for Concept Maps	48	25
	23	

Aluno Teta

IDENTIFICAÇÃO	ζ zeta			ANÁLISE FERRAMENTA			ANÁLISE ESPECIALISTA		
	NÍVEL	CONCEITO 1	VERBO DE LIGAÇÃO	CONCEITO 2	ESTRUTURA	GRAMÁTICA	CONTEXTO	ESTRUTURA	GRAMÁTICA
N0	PLN	Busca solucionar	Questões computacionais	SIM	SIM	SIM	SIM	SIM	SIM
N0	PLN	Fornecer aos computadores	Capacidade entender textos	SIM	SIM	SIM	SIM	SIM	SIM
N0	PLN	Sub-área da	Inteligência artificial	SIM	NÃO		SIM	NÃO	
N1	Questões computacionais	Que inclui	Capacidade entender textos	SIM	SIM	SIM	SIM	SIM	SIM
N1	Capacidade entender textos	Criar	Resumos	SIM	SIM	SIM	SIM	SIM	SIM
N1	Capacidade entender textos	Extraí	Informações	SIM	SIM	NÃO	SIM	SIM	SIM
N1	Capacidade entender textos	Interpretar	Sentidos	SIM	SIM	SIM	SIM	SIM	SIM
N1	Capacidade entender textos	Necessidade de	Pré-processamentos	SIM	NÃO		SIM	SIM	
N1	Capacidade entender textos	Que significa analisar	Léxicamente	SIM	SIM	SIM	SIM	SIM	SIM
N1	Capacidade entender textos	Que significa analisar	Semânticamente	SIM	SIM	SIM	SIM	SIM	SIM
N1	Capacidade entender textos	Que significa analisar	Sentimentos	SIM	SIM	SIM	SIM	SIM	SIM
N1	Capacidade entender textos	Que significa analisar	Sintaticamente	SIM	SIM	SIM	SIM	SIM	SIM
N2	Capacidade entender textos	Criar	Resumos	SIM	SIM	SIM	SIM	SIM	SIM
N2	Capacidade entender textos	Extraí	Informações	SIM	SIM	NÃO	SIM	SIM	SIM
N2	Capacidade entender textos	Interpretar	Sentidos	SIM	SIM	SIM	SIM	SIM	SIM
N2	Capacidade entender textos	Necessidade de	Pré-processamentos	SIM	NÃO		SIM	SIM	
N2	Capacidade entender textos	Que significa analisar	Léxicamente	SIM	SIM	SIM	SIM	SIM	SIM
N2	Capacidade entender textos	Que significa analisar	Semânticamente	SIM	SIM	SIM	SIM	SIM	SIM
N2	Capacidade entender textos	Que significa analisar	Sentimentos	SIM	SIM	SIM	SIM	SIM	SIM
N2	Capacidade entender textos	Que significa analisar	Sintaticamente	SIM	SIM	SIM	SIM	SIM	SIM
N2	Pré-processamentos	Deixando apenas	Informações relevantes	SIM	SIM	NÃO	SIM	SIM	NÃO
N3	Pré-processamentos	Deixando apenas	Informações relevantes	SIM	SIM	NÃO	SIM	SIM	NÃO
							22	20	17
							100%	91%	89%
			ANÁLISE FERRAMENTA			ANÁLISE ESPECIALISTA			

Scoring Criteria for Concept Maps	64	28
	36	

Aluno Zeta

APÊNDICE D – VALIDAÇÕES DOS SCORES

```

-----  ::Scores para Proposições Válidas::  -----
5 proposicoes
relacionamentos * 3 = 15
maior hierarquia * 5 = 10
score de ramificacoes: 5
score positivo = 30

-----

-----  ::Scores para Proposições Inválidas::  -----
5 proposicoes
relacionamentos * 3 = 15
maior hierarquia * 5 = 5
Ramificações
score de ramificacoes: 5
score negativo = 25

-----

----- CROSS -----

[]

-----

----- CROSS -----

Score Total 5

```

Aluno Beta

```

-----  ::Scores para Proposições Válidas::  -----
9 proposicoes
relacionamentos * 3 = 27
maior hierarquia * 5 = 10
score de ramificacoes: 9
score positivo = 46

-----

-----  ::Scores para Proposições Inválidas::  -----
1 proposicoes
relacionamentos * 3 = 3
maior hierarquia * 5 = 10
score de ramificacoes: 4
score negativo = 17

-----

----- CROSS -----

[]

-----

----- CROSS -----

Score Total 29

```

Aluno Delta

```

-----  ::Scores para Proposições Válidas::  -----

4 proposicoes
relacionamentos * 3 = 12
maior hierarquia * 5 = 15
score de ramificacoes: 0
score positivo = 27

-----

-----  ::Scores para Proposições Inválidas::  -----

1 proposicoes
relacionamentos * 3 = 3
maior hierarquia * 5 = 5
score de ramificacoes: 0
score negativo = 8

-----

----- CROSS -----

[]

-----

----- CROSS -----

Score Total 19
    
```

Aluno Epsilon

```

-----  ::Scores para Proposições Válidas::  -----

15 proposicoes
relacionamentos * 3 = 45
maior hierarquia * 5 = 15
score de ramificacoes: 5
score positivo = 65

-----

-----  ::Scores para Proposições Inválidas::  -----

0 proposicoes
relacionamentos * 3 = 0
maior hierarquia * 5 = 0
score de ramificacoes: 0
score negativo = 0

-----

----- CROSS -----

[]

-----

----- CROSS -----

Score Total 65
    
```

Aluno Gama

```

----- .:Scores para Proposições Válidas:. -----
14 proposicoes
relacionamentos * 3 = 42
maior hierarquia * 5 = 30
score de ramificacoes: 5
score positivo = 77

-----

----- .:Scores para Proposições Inválidas:. -----
0 proposicoes
relacionamentos * 3 = 0
maior hierarquia * 5 = 0
score de ramificacoes: 0
score negativo = 0

-----

----- CROSS -----

[]

-----

----- CROSS -----

Score Total 77
    
```

Aluno Lambda

```

----- .:Scores para Proposições Válidas:. -----
2 proposicoes
relacionamentos * 3 = 6
maior hierarquia * 5 = 5
score de ramificacoes: 0
score positivo = 11

-----

----- .:Scores para Proposições Inválidas:. -----
1 proposicoes
relacionamentos * 3 = 3
maior hierarquia * 5 = 0
score de ramificacoes: 0
score negativo = 3

-----

----- CROSS -----

[]

-----

----- CROSS -----

Score Total 8
    
```

Aluno Sigma

```

----- .:Scores para Proposições Válidas:. -----
8 proposicoes
relacionamentos * 3 = 24
maior hierarquia * 5 = 20
score de ramificacoes: 4
score positivo = 48

-----

----- .:Scores para Proposições Inválidas:. -----
3 proposicoes
relacionamentos * 3 = 9
maior hierarquia * 5 = 10
score de ramificacoes: 4
score negativo = 23

-----

----- CROSS -----

[]

-----

----- CROSS -----

Score Total 25

```

Aluno Teta

```

----- .:Scores para Proposições Válidas:. -----
15 proposicoes
relacionamentos * 3 = 45
maior hierarquia * 5 = 10
score de ramificacoes: 9
score positivo = 64

-----

----- .:Scores para Proposições Inválidas:. -----
4 proposicoes
relacionamentos * 3 = 12
maior hierarquia * 5 = 15
score de ramificacoes: 9
score negativo = 36

-----

----- CROSS -----

[]

-----

----- CROSS -----

Score Total 28

```

Aluno Zeta

APÊNDICE E – CLASSIFICAÇÃO DA TAXONOMIA TOPOLÓGICA

Nível 4
Sem explicações longas: 41
Não faltam Palavras de Link (Verbo de ligação): 41
Ramificação Alta (5-6 pontos de ramificação): 5
3 ou mais níveis de hierarquia: 3

Aluno Beta

.: Sem Categoria :.
.: Entre N2 e N3 :.
Sem explicações longas: 19
Não faltam Palavras de Link (Verbo de ligação): 19
Ramificação Média: 3
Níveis de hierarquia: 3

Aluno Delta

.: Sem Categoria :.
.: Entre N2 e N3 :.
Sem explicações longas: 10
Não faltam Palavras de Link (Verbo de ligação): 10
Ramificação Média: 2
Níveis de hierarquia: 3

Aluno Epsilon

Nível 4
Sem explicações longas: 34
Não faltam Palavras de Link (Verbo de ligação): 34
Ramificação Alta (5-6 pontos de ramificação): 5
3 ou mais níveis de hierarquia: 3

Aluno Gama

Nível 4
Sem explicações longas: 21
Não faltam Palavras de Link (Verbo de ligação): 21
Ramificação Alta (5-6 pontos de ramificação): 6
3 ou mais níveis de hierarquia: 6

Aluno Lambda

```
.: Sem Categoria :.  
.: Entre N2 e N3 :.  
Sem explicações longas: 16  
Não faltam Palavras de Link (Verbo de ligação): 16  
Ramificação Média: 3  
Níveis de hierarquia: 3
```

Aluno Zeta

Os Alunos SIGMA e TETA não obtiveram classificação

Anexos

ANEXO A – SCORING CRITERIA FOR CONCEPT MAPS

Protocolo de pontuação para o *Scoring Criteria for Concept Maps*:

Categoria	Valid	Not Valid (-)	Toral Valid Score	Comments
Relationships to each link 1 = not labeled 1 _____ = _____ (not labeled)				
3 = labeled 3 _____ = _____ (labeled)				
Hierarchy 5 points each level 5 _____ = _____ (# of Levels)				
Branching Level 1 = 5 Points 5 _____ = _____ (# of Branches)				
Level 2 = 4 Points 4 _____ = _____ (# of Branches)				
Level 3 = 3 Points 3 _____ = _____ (# of Branches)				
Level 4 = 2 Points 2 _____ = _____ (# of Branches)				
Level 5 & beyond = 1 Point 1 _____ = _____ (# of Branches)				
cross links 10 points each 10 _____ = _____ (# of cross-links)				
Examples 1 point each 1 _____ = _____ (# of examples)				
Nonexamples 1 point each 1 _____ = _____ (# of examples)				
Total				

ANEXO B – TAXONOMIA TOPOLÓGICA

Taxonomía Topológica de Cañas et al. (2006)

Nivel 0

- a) *Predominan explicaciones largas sobre conceptos*
- b) *Sin palabras de enlace*
- c) *Lineal (0-1 puntos de ramificación)*

Nivel 1

- a) *Predominan conceptos sobre explicaciones largas*
- b) *Faltan la mitad o más de las palabras de enlace*
- c) *Lineal (0-1 puntos de ramificación)*

Nivel 2

- a) *Predominan conceptos sobre explicaciones largas*
- b) *Faltan menos de la mitad de las palabras de enlace*
- c) *Ramificación baja (2 puntos de ramificación)*

Nivel 3

- a) *Sin explicaciones largas*
- b) *No faltan palabras de enlace*
- c) *Ramificación media (3-4 puntos de ramificación)*
- d) *Menos de 3 niveles de jerarquía*

Nivel 4

- a) *Sin explicaciones largas*
- b) *No faltan palabras de enlace*
- c) *Ramificación alta (5-6 puntos de ramificación)*
- d) *3 o más niveles de jerarquía*

Nivel 5

- a) *Sin explicaciones largas*
- b) *No faltan palabras de enlace*
- c) *Ramificación alta (5-6 puntos de ramificación)*
- d) *3 o más niveles de jerarquía*
- e) *De 1-2 enlaces cruzados*

Nivel 6

- a) *Sin explicaciones largas*
- b) *No faltan palabras de enlace*
- c) *Ramificación muy alta (7 o más puntos de ramificación)*
- d) *3 o más niveles de jerarquía*
- e) *Más de 2 enlaces cruzados*