

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Ricardo Radaelli Meira

**Ferramenta computacional para avaliação automatizada da escrita
de projetos de pesquisa**

Porto Alegre

1. Semestre

2024

Ricardo Radaelli Meira

**Ferramenta computacional para avaliação automatizada da escrita
de projetos de pesquisa**

Tese do Curso de Pós-Graduação em Informática na Educação, da Universidade Federal do Rio Grande do Sul, como requisito para a obtenção do Título de Doutor(a) em Informática na Educação.

Orientador(a): Prof. Dr. Eliseo Berni Reategui

Porto Alegre

1. Semestre

2024

CIP - Catalogação na Publicação

Meira, Ricardo Radaelli
Ferramenta computacional para avaliação
automatizada da escrita de projetos de pesquisa /
Ricardo Radaelli Meira. -- 2024.
127 f.
Orientador: Eliseo Berni Reategui.

Tese (Doutorado) -- Universidade Federal do Rio
Grande do Sul, Centro de Estudos Interdisciplinares em
Novas Tecnologias na Educação, Programa de
Pós-Graduação em Informática na Educação, Porto
Alegre, BR-RS, 2024.

1. Feedback automatizado. 2. Avaliação automatizada
da escrita. 3. Analítica da escrita. 4. Projetos de
pesquisa. I. Reategui, Eliseo Berni, orient. II.
Titulo.



UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
CINTED – CENTRO INTERDISCIPLINAR DE NOVAS TECNOLOGIAS NA EDUCAÇÃO
PPGIE – PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA NA EDUCAÇÃO

**ATA SOBRE A DEFESA DE TESE DE DOUTORADO
RICARDO RADAELLI MEIRA**

Às quatorze horas do dia quinze de julho de dois mil e vinte e quatro, no endereço eletrônico <https://mconf.ufrgs.br/webconf/00126356>, conforme a portaria 02 de 10/10/2022 da PROPG/UFRGS que regulamenta a modalidade híbrida ou a distância para as bancas de defesas de cursos *stricto sensu*, reuniu-se a Comissão de Avaliação, composta pelos Professores Doutores: Prof.^a Dr.^a Patrícia da Silva Campelo Costa Barcellos (PPGIE/UFRGS), Prof.^a Dr.^a Elisa Boff (UCS), Prof.^a Dr.^a Karla Marques da Rocha (UFSM), para a análise da Defesa de Tese de Doutorado intitulada *Ferramenta Computacional para Avaliação Automatizada da Escrita de Projetos de Pesquisa*, do doutorando de Pós-Graduação em Informática na Educação Ricardo Radaelli Meira sob a orientação do Prof. Dr. Eliseo Berni Reategui. A Banca, reunida, após a apresentação e arguição, emite o parecer abaixo assinalado.

[x] Considera a Tese Aprovada

() sem alterações;

(x) sem alterações, com voto de louvor;

() e recomenda que sejam efetuadas as reformulações e atendidas as sugestões contidas nos pareceres individuais dos membros da Banca;

[] Considera a Tese Reprovada.

Considerações adicionais (a critério da Banca):

Documento assinado digitalmente

gov.br

ELISEO BERNI REATEGUI

Data: 15/07/2024 16:54:38-0300

Verifique em <https://validar.iti.gov.br>

Prof. Dr. Eliseo Berni Reategui

Orientador

Documento assinado digitalmente

gov.br

PATRICIA DA SILVA CAMPELO COSTA BARCELL

Data: 16/07/2024 14:17:20-0300

Verifique em <https://validar.iti.gov.br>

Prof.^a Dr.^a Patrícia da Silva Campelo Costa

Barcellos

PPGIE/UFRGS

Documento assinado digitalmente

gov.br

ELISA BOFF

Data: 15/07/2024 21:42:20-0300

Verifique em <https://validar.iti.gov.br>

Prof.^a Dr.^a Elisa Boff

UCS

Documento assinado digitalmente

gov.br

KARLA MARQUES DA ROCHA

Data: 17/07/2024 08:53:50-0300

Verifique em <https://validar.iti.gov.br>

Prof.^a Dr.^a Karla Marques da Rocha

UFSM

RESUMO

A capacidade de se comunicar de forma escrita é de extrema importância para uma vida plena em sociedade. No meio acadêmico, essa habilidade se torna ainda mais importante. A escrita acadêmica, com todas as suas particularidades, demanda tanto do escritor como do avaliador, o domínio sobre diversos conceitos e normas acadêmicas. Uma das formas de diminuir o peso dessa tarefa consiste no uso da Analítica da Aprendizagem, de técnicas de Processamento de Linguagem Natural e de ferramentas de avaliação por parte do professor. Dessa forma, este trabalho procura investigar como uma ferramenta computacional com funções de avaliação automatizada da escrita pode auxiliar o professor na avaliação e elaboração de feedback personalizado para os alunos sobre seus textos do gênero projeto de pesquisa. Para isso, realizou-se um experimento com professores de cursos de graduação de disciplinas de metodologia científica ou da pesquisa onde foi aplicada uma ferramenta de avaliação automatizada de projetos de pesquisa que forneceu feedbacks a esses profissionais, buscando auxiliar na tarefa de avaliação. Como metodologia deste estudo, utilizou-se o estudo de caso múltiplos e os instrumentos de coleta de dados foram entrevistas e dados de uso da ferramenta desenvolvida. Os resultados obtidos mostram a capacidade da ferramenta em avaliar projetos de pesquisa e fornecer feedbacks que auxiliam os professores e também na identificação da aceitação, tanto da ferramenta quanto dos feedbacks gerados.

Palavras-chave: Feedback automatizado; Avaliação automatizada da escrita; Projetos de pesquisa.

ABSTRACT

The ability to communicate effectively in writing is of utmost importance for a fulfilling life in society. In the academic realm, this skill becomes even more crucial. Academic writing, with all its nuances, demands mastery of various concepts and academic norms from both the writer and the evaluator. One way to alleviate the burden of this task is through the use of Learning Analytics, Natural Language Processing techniques, and evaluation tools by educators. Consequently, this study aims to investigate how a computational tool with automated writing assessment capabilities can assist teachers in evaluating and providing personalized feedback to students on their research project proposals. To this end, an experiment will be conducted with university professors teaching scientific methodology or research courses. They will use an automated assessment tool for research projects that provides feedback, aiming to aid in the evaluation process. The methodology of this study is a case study, and data collection tools will include interviews, questionnaires, and usage data from the developed tool. The results demonstrate the tool's ability to evaluate research projects and provide feedback that assists teachers, as well as revealing the acceptance of both the tool and the generated feedback.

Keywords: Automated feedback; Automated writing assessment; Research project.

LISTA DE FIGURAS

Figura 1 - Gráfico de resultado do SOBEK	46
Figura 2 - Campos do plug-in MTA	52
Figura 3 - Fluxograma do plug-in MTA	53
Figura 4 - Extração de conceitos chave plug-in MTA	54
Figura 5 - Exemplo de texto avaliado pelo MTA	55
Figura 6 - Feedback do plug-in MTA	56
Figura 7 - Diagrama de funcionamento da ferramenta Athos	60
Figura 8 - Gráfico da análise estrutural dos projetos	61
Figura 9 - Gráfico de complexidade textual dos projetos	62
Figura 10 - Tela de atividades enviadas pelo aluno.	63
Figura 11 - Tela de atividades do professor.	64
Figura 12 - Tela de validação de extrações.	65
Figura 13 - Tela de avaliação e feedback.	66
Figura 14 - Disposição de z-scores por conceito	76
Figura 15 - Gráfico de ocorrência das métricas de ET - Projetos	81
Figura 16 - Disposição dos feedbacks por métricas	89
Figura 17 - Comentário sobre o feedback gerado	94
Figura 18 - Comentário do professor participante 2 - Métrica "QuantAções"	95
Figura 19 - Comentário professor participante 1 - Métrica "QuantAções"	96
Figura 20 - Comentário sobre apresentação do objetivo.	98
Figura 21 - Quantidade de ocorrência dos códigos	100
Figura 22 - Comentário sobre a importância da extração dos dados	102
Figura 23 - Comentário sobre alterações no modo da escrita	103

LISTA DE TABELAS

Tabela 1 - Apresentação das métricas da ferramenta Athos	68
Tabela 2 - Tabela de dados ANOVA Notas x Métrica	72
Tabela 3 - Quadro de resultados para o corpus Projetos de Pesquisa	75
Tabela 4 - Tabela de correlações para notas e métricas	78
Tabela 5 - Ocorrência das Métricas de Estrutura Textual	80
Tabela 6 - Correlação de Pearson para métricas de Estrutura Textual	82
Tabela 7 - Taxa de aceitação dos feedbacks	91
Tabela 8 - Resultados da extração automatizada	97
Tabela 9 - Códigos da Análise de Conteúdo	99

SUMÁRIO

1 INTRODUÇÃO	6
2 REFERENCIAL TEÓRICO	11
2.1 O ENSINO DA ESCRITA	11
2.1.1 O professor como mediador	13
2.1.2 O professor e o ensino da escrita de trabalhos acadêmicos	16
2.1.3 O feedback do professor mediador	18
2.1.4 Feedback automatizado e feedback personalizado	22
2.2 A ANALÍTICA DA APRENDIZAGEM	24
2.2.1 Analítica da escrita	26
2.2.2 Processamento de Linguagem Natural (PLN)	28
2.2.2.1 Ferramentas de PLN	29
2.2.2.2 Tokenizadores	31
2.2.2.3 Parsers	33
2.2.2.4 Lematizadores	35
2.3 A AVALIAÇÃO DE PROJETOS DE PESQUISA	36
2.3.1 Critérios de avaliação de projetos de pesquisa	37
2.3.2 Avaliação automática de projetos de pesquisa	40
2.4 TECNOLOGIAS PARA ANÁLISE DA ESCRITA	42
2.4.1 Coh-Metrix e NILC-Metrix	43
2.4.2 Sobek	45
2.4.3 Ferramentas de feedback automatizado da escrita	46
2.4.4 O plug-in MTA	51
3 METODOLOGIA DA PESQUISA E RESULTADOS	57
3.1 Etapa 1: Desenvolvimento da ferramenta	58
3.1.2. Resultados sobre o desenvolvimento da ferramenta	67
3.2 Etapa 2: Análise de métricas de complexidade e estrutura textual	70
3.2.1 Metodologia para a Análise de Redações seguindo critérios do ENEM	70
3.2.2 Resultados para a Análise de Redações seguindo critérios do ENEM	71
3.2.3 Metodologia para a Análise dos Projetos de Pesquisa de Graduação	75
3.2.4 Resultados para a Análise dos Projetos de Pesquisa de Graduação	76
3.3 Etapa 3: Estudo Final para Avaliação do Emprego da ferramenta Athos	84
3.3.1 Metodologia do Estudo Final para Avaliação do Emprego da ferramenta	85
3.4.2. Resultados do Estudo Final para Avaliação do Emprego da ferramenta	89
4 CONCLUSÕES	106
4.1 Considerações Finais	108
REFERÊNCIAS	111

1 INTRODUÇÃO

Escrever é um ato comum e essencial ao ambiente escolar, desde o ensino fundamental até a pós-graduação, o estudante se depara com situações onde o ato de escrever é indispensável para a resolução do desafio, independentemente do nível de ensino em que se encontre. Conforme esse aluno avança e se desenvolve academicamente, novas habilidades de escrita são exigidas. No ensino superior, o aluno é incentivado a escrever de acordo com os níveis de exigências dos chamados textos acadêmicos, atingindo um novo padrão de escrita.

Além da academia, a vida cotidiana também exige do indivíduo a habilidade de comunicação escrita de forma coerente. Através dessa habilidade, é possível ao homem em sociedade se expressar, demonstrar sentimentos, opiniões, desejos e diversos sentimentos que o façam interagir e permitam viver em sociedade (Weston-Sementelli; Alle; McNamara, 2018).

Ao longo das últimas décadas, uma série de novas ferramentas tecnológicas introduziram novas formas de escrita, revisão, análise, avaliação e compartilhamento. Nesse contexto, a difusão de novas tecnologias tornam o ensino e a aprendizagem da escrita cada vez mais importantes. Contudo, a mera existência dessas tecnologias não é capaz de trazer resultados positivos. Seu emprego cuidadoso é necessário para melhorar a eficiência instrucional de pedagogias e abordagens do ensino da escrita que possam enfatizar propósitos sociais significativos, domínio de gêneros relevantes e desenvolvimento da proficiência linguística acadêmica dos alunos (Warschauer, 2010).

Com o avanço da tecnologia educacional, especialmente nos campos do processamento de linguagem natural e da inteligência artificial, diversas ferramentas foram desenvolvidas para auxílio de alunos e professores no processo de ensino-aprendizagem da escrita. Dentre os diversos usos, destacam-se os trabalhos de Crossley et. al. (2010), Aluísio et. al. (2010), Xia et. al. (2016) e Pires et. al. (2017), os quais fizeram uso das referidas tecnologias com o objetivo de simplificação de textos ou da identificação de métricas textuais avaliadoras.

Juntamente com esse desenvolvimento, também surgiram preocupações e críticas em relação à falta de confiabilidade e à inadequação de uma máquina na correção da escrita dos alunos. Por exemplo, o Conselho Nacional de Professores de Inglês (NCTE) emitiu uma forte declaração contra o uso da pontuação por máquinas na educação em 2013. A declaração destaca que *"os computadores são*

incapazes de reconhecer ou julgar os elementos que mais associamos à boa escrita (lógica, clareza, precisão, ideias relevantes para um tópico específico, estilo inovador, apelos eficazes ao público, diferentes formas de organização, tipos de persuasão, qualidade das evidências, humor ou ironia e usos eficazes da repetição, para citar apenas alguns)" (NCTE, 2013). Assim, nos encontramos em um período em que começam a proliferar ferramentas tecnológicas de análise e avaliação automatizada de textos, tornando-se fundamental discutir e melhor compreender o enquadramento dos professores nesse cenário.

A avaliação automatizada da escrita tem sido usada para fornecer aos alunos feedback sobre suas produções textuais. Muitos estudos empíricos examinaram a eficácia desse tipo de estratégia no que diz respeito à melhoria da qualidade da escrita. Na revisão de literatura apresentada por Zhai e Ma (2022), os autores examinaram um conjunto de 26 estudos primários entre 2010 e 2022 revelando que a avaliação automatizada da escrita teve nessas pesquisas um efeito geral positivo, sobretudo no que diz respeito à escrita em língua estrangeira. Quanto aos gêneros textuais, os resultados mais significativos da Avaliação Automatizada da Escrita foram alcançados em atividades de escrita argumentativa.

A escrita acadêmica possui particularidades que muitas vezes não são conhecidas pelos alunos ingressantes na graduação. Para a escrita de um projeto de pesquisa, por exemplo, o aluno precisa conhecer e utilizar conceitos e normas acadêmicas, além de organizar, gerenciar e estruturar a ampla gama de informações que fundamentam seus textos. Deste modo, apenas conhecer a linguagem e suas regras não basta para a escrita de um bom texto, sendo comum que os alunos recém-chegados à graduação apresentem dificuldades na escrita de textos acadêmicos.

Na escrita de projetos, como no caso de trabalhos de conclusão de curso, Vergara (1998), cita os objetivos da pesquisa, o referencial teórico e a metodologia como componentes essenciais que precisam ser utilizados de acordo com normas científicas da comunidade acadêmica. Porém, se escritores experientes julgam não trivial a tarefa de escrever esse texto, indivíduos que estão sendo introduzidos aos conceitos desse gênero de escrita têm ainda maior dificuldade, uma vez que não estão familiarizados com suas diversas convenções e processos (Assis, 2014).

Em nosso sistema educacional, uma atividade é comumente avaliada por um instrutor competente, capaz de corrigir erros e orientar os alunos para que possam melhorar em atividades futuras. No ambiente educacional brasileiro, onde muitas vezes o professor tem muitas turmas e/ou turmas com muitos alunos, fazer diversas atividades de escrita e as correções que estas demandam, exigem do profissional um longo tempo de preparo e finalização, o que pode reduzir a frequência com que essas atividades são solicitadas aos alunos. Em disciplinas de produção científica, comuns em programas de graduação e pós-graduação, essa tarefa de correção se torna ainda mais laboriosa, uma vez que os textos aumentam em complexidade, gama de assuntos e tamanho (Araújo et al., 2017).

Nesse contexto, a avaliação automatizada da escrita poderia entrar em cena. Porém, para os professores, ferramentas de feedback automático podem se mostrar como “caixas-pretas” que não são transparentes em seu modo de operação e que também não foram desenhadas com um papel claro para o professor no processo de ensino e aprendizagem. Timms (2016) argumenta que, nos próximos 25 anos, deverão ser produzidos sistemas que operem como “co-bots educacionais”, ou seja, sistemas que atuem em colaboração com o professor para dar um apoio mais personalizado aos estudantes. É baseado nessa ideia que este projeto foi desenvolvido.

Garcia e Benevides-Pereira (2005) salientam que raramente os professores dispõem de tempo suficiente para ler, editar e orientar os alunos em seus projetos, o que faz com que estes alunos, muitas vezes, trabalhem de forma autônoma sem o devido conhecimento para tal. Deste modo, torna-se relevante o desenvolvimento de ferramentas que possam apoiar o trabalho de mediação do professor no ensino e aprendizagem da escrita.

Neste contexto, foi definido o objetivo geral desta tese: *Investigar como uma ferramenta computacional com funções de avaliação automatizada da escrita pode auxiliar o professor na avaliação e elaboração de feedback personalizado para os alunos sobre seus textos do gênero projeto de pesquisa.*

Os objetivos específicos, definidos em torno do objetivo geral, são:

- Elencar funções de avaliação automatizada da escrita que possam ser utilizadas para auxiliar o professor na avaliação de textos acadêmicos;

- Projetar um modelo de sistema para a avaliação da escrita de trabalhos acadêmicos que permita ao professor acompanhar e dar feedback personalizado aos alunos sobre suas produções;
- Implementar e avaliar o modelo desenvolvido em contexto real de ensino e aprendizagem no ensino superior.

Assim, a questão norteadora desta pesquisa é a seguinte: como uma ferramenta computacional com funções de avaliação automatizada da escrita pode auxiliar o professor na avaliação e elaboração de feedback personalizado para os alunos sobre seus textos do gênero projeto de pesquisa?

Na seleção de técnicas e ferramentas de avaliação automatizada de produções textuais, serão consideradas funções de análise de texto utilizadas no processamento de linguagem natural, bem como ferramentas de mineração de texto. Neste contexto, a pesquisa se enquadra também na discussão sobre o emprego da Inteligência Artificial (IA) na Educação, seu potencial e desafios. Como afirmam Luckin et al. (2016), caberá aos professores compreender e orquestrar sobre quando e como usar as ferramentas de Inteligência Artificial em suas práticas. Na tese aqui apresentada, o papel do professor foi pensado desde o desenho da ferramenta tecnológica, já que ela foi desenvolvida tendo o professor como ator central na mediação do processo de ensino e aprendizagem da escrita acadêmica.

A pesquisa foi desenvolvida na modalidade de estudo de caso descritivo, buscando compreender e descrever *como os docentes fazem uso de uma ferramenta com funções de avaliação da escrita para auxiliá-los na avaliação e elaboração de feedback personalizado para os alunos*. O estudo envolveu o acompanhamento de 2 professores de disciplinas de metodologia de pesquisa em Educação, buscando registrar o modo como empregam as ferramentas desenvolvidas na produção de feedback para os alunos bem como no registro de suas percepções sobre a utilidade e interesse da ferramenta. A pesquisa seguirá abordagem quali-quantitativa.

No que diz respeito à originalidade da pesquisa, esta tese contribui com o avanço do conhecimento na área em diferentes aspectos:

- identificação e validação de funções para avaliação de projetos de pesquisa, ampliando conjunto de funções desenvolvido em pesquisa anterior (Bastiani, 2022);
- apoio à avaliação da escrita de projetos com foco na elaboração de feedback para os alunos com a intervenção do professor.
- identificação e definição de métricas de avaliação textual capazes de auxiliar o professor na identificação e correção de trabalhos.

Apesar de este projeto ter como foco a escrita acadêmica, as funções de avaliação consideradas na pesquisa deverão trazer resultados que devem poder ser aplicados na avaliação da produção de outros gêneros textuais. Neste sentido, cabe ressaltar a relevância da avaliação de produções escritas em um cenário em que cada vez mais testes de múltipla escolha são eleitos como principal forma de avaliação por sua facilidade de aplicação e correção, mesmo em contextos online.

Este trabalho está segmentado em quatro partes. A primeira, chamada capítulo 2, apresenta o embasamento teórico através de uma revisão bibliográfica sobre o ensino da escrita, o papel do professor mediador, a analítica da aprendizagem, o processamento de linguagem natural e as ferramentas tecnológicas para a avaliação da escrita. O capítulo 3 contém os métodos utilizados na pesquisa e as formas e bases em que o estudo foi realizado. A última parte do documento contém os resultados, conclusões e discussões.

2 REFERENCIAL TEÓRICO

Nesta seção do trabalho, é apresentado o embasamento teórico da pesquisa, este embasamento está dividido em quatro subseções, sendo a seção 2.1 responsável por abranger o feedback no ensino da escrita, a seção 2.2, aborda a analítica da aprendizagem, a seção 2.3 apresenta a avaliação de projetos de pesquisa e a seção 2.4 traz as tecnologias para análise e avaliação da escrita.

2.1 O ENSINO DA ESCRITA

A escrita é uma das habilidades fundamentais para a comunicação humana e um dos pilares básicos da educação formal. No entanto, o ensino da escrita ainda é um desafio para muitos educadores, especialmente quando se trata de lidar com a diversidade de habilidades, experiências e expectativas dos alunos. Nesse contexto, é importante considerar as abordagens pedagógicas que podem ajudar a aprimorar o ensino da escrita e, conseqüentemente, o desenvolvimento da competência escrita dos alunos.

Segundo Lourenço e Paiva (2010), uma das maiores dificuldades no ensino da escrita é a falta de motivação dos estudantes. O autor afirma que muitas vezes os alunos não enxergam a importância da escrita em suas vidas, o que acaba tornando a aprendizagem mais difícil. Além disso, muitos estudantes têm dificuldades em estruturar suas ideias e em organizar o texto de forma coerente e coesa.

Outro problema apontado por Pereira (2001) é a falta de incentivo à leitura por parte dos pais e educadores. Segundo as autoras, a leitura é uma das principais ferramentas para o desenvolvimento da escrita, pois permite que o aluno tenha contato com diferentes gêneros textuais e amplie seu vocabulário. Quando os alunos não têm o hábito de ler, a escrita pode se tornar uma tarefa difícil e desafiadora.

Além disso, é importante considerar as diferenças individuais dos alunos, pois cada um possui um ritmo de aprendizagem e uma forma de aplicar o conhecimento. Por isso, é fundamental que os educadores adaptem suas

estratégias de ensino às necessidades de cada aluno, para garantir que todos possam desenvolver suas habilidades de escrita da melhor forma possível.

Para superar essas dificuldades, é preciso adotar estratégias que motivem os estudantes e os incentivem a desenvolver suas habilidades de escrita. Segundo Pereira (2001), é importante criar atividades que despertem o interesse dos alunos, como produção de textos sobre temas atuais e relevantes. Além disso, é fundamental que os educadores incentivem a leitura e ofereçam feedbacks construtivos para que os alunos possam aprimorar suas habilidades de escrita.

Dessa forma, todo ensino da escrita deve se orientar pelos objetivos do gênero textual a ser desenvolvido. Na escrita acadêmica, foco deste trabalho, o professor deve orientar seus alunos para as convenções e normas que esse tipo de trabalho preconiza. Assim, o autor conseguirá desenvolver a forma de comunicação adaptada ao conhecimento produzido para seus pares e a comunidade científica (Marcuschi, 2010).

Uma das teorias que pode ser utilizada no ensino da escrita acadêmica é a abordagem sociointeracionista, que destaca a importância do diálogo e da interação social no processo de ensino e aprendizagem da escrita. Conforme Vygotsky (1991), o aprendizado ocorre por meio da interação com o outro, o que inclui a troca de experiências e o feedback constante. A partir dessa abordagem pode-se valorizar a produção textual em sala de aula e a interação entre os alunos, que podem se ajudar mutuamente a desenvolver suas habilidades de escrita.

Além da troca de conhecimentos e produções em sala de aula entre os alunos, cabe aqui destacar a importância do professor como mediador do conhecimento a ser construído (Perrenoud, 1999).

Vygotsky (1991) fala da importância do feedback constante, seja pelo mediador ou por meio da interação social com os pares. No ensino da escrita, isto deve ser considerado. Um método frequente no ensino da escrita é aquele em que o professor é tomado como modelo de escritor. Os alunos observam e procuram seguir seus ensinamentos, buscam reproduzir os exemplos que fornece de produção textual. Lerner (2002), corrobora com essa ideia argumentando que o ensino da

escrita deve levar em conta a importância do exemplo, já que os alunos tendem a se inspirar nos modelos de escrita que observam em seu ambiente.

Assim, o papel do professor como mediador no ensino da escrita envolve que esse seja capaz de identificar as necessidades individuais dos alunos, criar um ambiente de aprendizagem favorável à interação social, disponibilizar exemplos de escrita que os alunos possam tomar como modelo e prover feedback adequado aos alunos quando esses realizam suas próprias produções textuais..

2.1.1 O professor como mediador

O ensino da escrita acadêmica é uma tarefa complexa que envolve a aquisição de habilidades específicas para a produção de textos de qualidade e coerentes com as normas acadêmicas. Nesse sentido, o papel do professor como mediador no processo de aprendizagem da escrita é fundamental, uma vez que cabe a ele fornecer orientações e feedback que possibilite aos estudantes desenvolverem suas habilidades de escrita e aprimorarem suas produções.

Diversas teorias apontam para o papel do professor como mediador no ensino da escrita acadêmica. Segundo Vygotsky (1998), o processo de aprendizagem ocorre a partir das interações sociais, e o professor, como mediador, deve ser capaz de fomentar o pensamento crítico e reflexivo dos alunos. Nessa perspectiva, o professor é visto como um mediador que atua como um guia para o aluno, auxiliando-o a construir o conhecimento e a desenvolver habilidades cognitivas por meio de interações sociais.

No contexto da escrita, o mediador pode ajudar o aluno a desenvolver sua capacidade de escrita, orientando-o em relação à organização das ideias, à escolha das palavras e à estruturação do texto. Ele pode também fornecer feedback construtivo sobre o trabalho do aluno, ajudando-o a identificar pontos fortes e áreas que precisam ser melhoradas.

Nesse sentido, é importante que o mediador tenha uma postura de facilitador, encorajando o aluno a se expressar e a explorar diferentes possibilidades de escrita.

No que diz respeito à avaliação do trabalho do aluno, Vygotsky (1998) destaca que o feedback deve ajudar o aluno a entender seus pontos fortes e fracos e a desenvolver sua capacidade de reflexão. O feedback fornecido pelo mediador experiente pode ajudar o aluno a desenvolver sua habilidade de escrita de forma mais eficaz, fornecendo orientação sobre como melhorar a estrutura, a organização e a coesão do texto. Segundo Hattie (2008), "o feedback eficaz é uma das estratégias mais poderosas para melhorar o aprendizado dos alunos, pois oferece informações específicas sobre o desempenho em relação a um objetivo e orienta os alunos sobre como aprimorar seu desempenho" (Hattie, 2008, p. 173).

Estudos como o de Carless e Boud (2018) têm destacado a importância do feedback no processo de aprendizagem, especialmente no contexto da escrita acadêmica ao afirmar que "o feedback é essencial para ajudar os alunos a melhorar sua capacidade de escrita e a desenvolver uma compreensão mais profunda dos requisitos da escrita acadêmica" (p. 19). Nesse contexto, os educadores apontam que um obstáculo importante para o sucesso da escrita é o domínio insuficiente da gramática lexical na composição do estilo acadêmico (Carless, 2007). Por outro lado, a tendência entre muitos professores de ver a gramática como uma forma separada de suas funções discursivas pode dificultar o processo de mediação do ensino da escrita. Neste sentido, apenas a correção de erros de gramática não é suficiente para tornar a escrita eficaz em contextos discursivos exigidos no ensino da escrita.

Entende-se assim, que o professor como mediador do processo e do ensino da escrita de projetos acadêmicos deve, além de apresentar as normas e modelos do gênero da escrita, identificar as limitações de seus alunos, ao mesmo tempo que busca promover a interação desses com o ambiente da escrita.

Nota-se assim a importância do papel mediador do professor em relação ao potencial de aprendizado e desenvolvimento do aluno. O professor é responsável por fornecer o suporte e a orientação necessários para que o aluno possa progredir na aprendizagem. Isso envolve a identificação dos objetivos de aprendizado, a seleção de atividades adequadas e a adaptação do ensino para atender às necessidades e habilidades individuais de cada aluno. O feedback dado pelo professor para as atividades realizadas pelos alunos cumpre esse papel,

ajudando-os a progredir. Ao fornecer comentários específicos sobre o trabalho dos alunos, os professores podem ajudá-los a entender melhor os conceitos e habilidades que estão tentando desenvolver e orientá-los na direção certa para alcançar suas metas educacionais. Esse tipo de feedback pode ser particularmente eficaz quando os professores identificam áreas específicas em que os alunos precisam melhorar e oferecem sugestões concretas e orientação para ajudá-los a alcançar suas metas. Ao receber feedback construtivo e personalizado, os alunos podem começar a desenvolver as habilidades e o conhecimento necessários para se tornarem mais independentes e bem-sucedidos em sua escrita.

Pode-se entender assim, que ferramentas de feedback podem auxiliar o desenvolvimento dos alunos. Dessa forma, softwares computacionais específicos para a geração de feedbacks automatizados têm sido desenvolvidos no contexto educacional. Contudo, para que sejam eficazes, esses feedbacks precisam ser específicos, descritivos e focados no processo de aprendizagem, em vez de apenas no produto final (HATTIE E TIMPERLEY, 2007). Nesse contexto, o feedback formativo, que olha para a "escrita futura do aluno" e visa auxiliá-lo em sua construção, tem sido preferido ao feedback somativo, que foca na escrita como um produto (HYLAND e HYLAND, 2006).

O feedback automatizado pode ter vantagens em relação ao feedback escrito do professor, ao dar um retorno imediato e personalizado ao aluno. Um estudo de Cavalcanti et. al. (2021), por exemplo, mostrou que o uso de feedbacks automatizados em um ambiente de aprendizado online pode melhorar o desempenho dos alunos e aumentar a motivação.

No entanto, projetar feedback automatizado que seja eficaz na identificação e apoio ao avanço individual do aluno é uma tarefa mais complexa. Os professores são formados para trabalhar nesse contexto, para identificar as dificuldades dos alunos e prover feedback/materiais de apoio que os auxiliem a avançar no desenvolvimento de suas habilidades e conhecimentos. É nessa direção que o feedback do professor pode tornar-se mais efetivo do que o feedback construído de maneira automática por uma máquina.

Além disso, Hyland e Hyland (2006) também observaram que o feedback por escrito do professor não é um elemento puramente informativo, pois para ser

eficaz deve engajar o autor. Pesquisas têm demonstrado que o feedback não é um componente descontextualizado, ao contrário, *"... é uma parte interativa de todo o contexto de aprendizagem, ajudando a criar um relacionamento interpessoal produtivo entre o professor e os alunos individualmente"* (p. 86). Nesse sentido, a própria organização das atividades de um curso deve garantir que os alunos possam interpretar e usar os comentários do professor de maneira eficaz.

Assim compreende-se a importância da mediação do professor e do feedback como ferramenta de auxílio ao aluno em sua evolução, no desenvolvimento de habilidades e aprendizagem.

2.1.2 O professor e o ensino da escrita de trabalhos acadêmicos

O ensino da escrita no ensino superior é uma questão complexa e desafiadora, que envolve muitos aspectos diferentes. A escrita é uma habilidade essencial para o sucesso acadêmico e profissional, e o ensino adequado da escrita é fundamental para garantir que os estudantes possam desenvolver essa habilidade de forma eficaz.

Pontos como a diversidade de habilidades e experiências dos alunos, a complexidade dos temas e assuntos acadêmicos, os objetivos da pesquisa e do currículo do curso são alguns dos desafios na tarefa do ensino da escrita de trabalhos acadêmicos.

Dentre as teorias que tratam do processo da escrita, algumas têm o foco no processo de tradução de ideias, planejamento e revisão do texto, como é a teoria de Hayes e Flower(1980). Porém, também existem pesquisadores que abordam o processo de escrita por outros ângulos, como é o caso dos trabalhos de Bazerman (2004), Hyland e Hyland (2006) e Hattie e Timperley (2007), enfatizando a importância do contexto social do escritor durante a tarefa de escrever. Bazerman (1981) possui trabalhos que debatem conceitos chamados de teoria sociocultural da escrita, esta teoria é baseada nos princípios da teoria sociocultural de Vygotsky (1978) e considerando a escrita como uma atividade socialmente mediada, influenciada por fatores contextuais, históricos e institucionais. Nesse caso, a teoria

considera como as práticas de escrita são moldadas pelas normas, valores e expectativas da comunidade. (Bazerman, 1981, 1997 e 2004)

A teoria sociocultural da escrita de Bazerman fundamenta-se na compreensão da escrita como um fenômeno profundamente enraizado nos contextos sociais e culturais em que ocorre. De acordo com Bazerman (2004), a escrita é um ato social e comunicativo que reflete as normas, valores e práticas compartilhadas por uma comunidade acadêmica ou disciplinar específica.

Nesse sentido, a escrita é influenciada por fatores socioculturais, como a estrutura organizacional das instituições de ensino, as expectativas dos professores e avaliadores, bem como as convenções disciplinares estabelecidas. A escrita acadêmica, por exemplo, exige a adoção de um estilo formal, o uso de citações e referências bibliográficas, e a conformidade com as normas de formatação estabelecidas pelas instituições (Swales e Feak, 2012).

Bereiter e Scardamalia (1987) reforçam que a escrita não é apenas uma forma de expressão de ideias preexistentes, mas um processo pelo qual são gerados e desenvolvidos novos significados e conhecimentos.

De acordo com essa perspectiva, a escrita desempenha um papel crucial na organização e na transformação do pensamento. Ao escrever, os estudantes são desafiados a articular seus conhecimentos, a refletir sobre suas ideias, a estabelecer conexões significativas entre os conceitos estudados e a transmitir isso para o futuro leitor. A escrita acadêmica, em particular, requer uma análise crítica e aprofundada de questões complexas, promovendo a construção de conhecimento mais sofisticado e aprofundado (Flower e Hayes, 1981).

A perspectiva sociocultural da escrita enfatiza que a escrita não é apenas uma atividade individual, mas uma prática que ocorre dentro de um sistema social mais amplo. Vygotsky (1978) argumenta que a linguagem escrita surge a partir de interações sociais e do desenvolvimento de habilidades cognitivas mediadas pela cultura. Segundo Vygotsky, a escrita é uma forma de linguagem secundária, que se diferencia da linguagem oral e permite a representação e a organização do pensamento de forma mais complexa.

Essa perspectiva sociocultural também destaca a importância das interações sociais na aprendizagem da escrita. Lave e Wenger (1991) ressaltam que a escrita acadêmica é um processo que se desenvolve em comunidades de prática, onde os estudantes interagem com seus pares, professores e especialistas para negociar significados e construir conhecimentos compartilhados.

Através do processo de escrita, os estudantes também são expostos a diferentes perspectivas e ideias. Ao pesquisar e interagir com textos acadêmicos, eles são desafiados a considerar diferentes pontos de vista, a avaliar evidências e a desenvolver argumentos convincentes. A escrita acadêmica, portanto, promove a capacidade dos estudantes de analisar criticamente informações, de pensar de forma independente e de contribuir para o avanço do conhecimento em suas áreas de estudo (Hyland e Hyland, 2006).

Além disso, a escrita também desempenha um papel essencial na consolidação do conhecimento. Escrever sobre um determinado tópico requer uma compreensão profunda dos conceitos e a habilidade de comunicar essas informações de maneira clara e coerente. Ao elaborar suas ideias por escrito, os estudantes são incentivados a organizar suas informações de forma lógica, a identificar lacunas em seu conhecimento e a buscar novas fontes de informação para aprimorar suas argumentações (Bazerman, 1997).

Assim, a escrita acadêmica não é apenas um produto final, mas um processo pelo qual o conhecimento é construído e desenvolvido. Através da reflexão, da análise crítica e da organização de ideias por escrito, os estudantes podem se engajar ativamente na construção de significados e na promoção de um pensamento mais complexo e sofisticado.

2.1.3 O feedback do professor mediador

Oferecer feedback na relação aluno-professor, como retorno de uma atividade ou correção traz ao estudante uma oportunidade de reflexão sobre suas habilidades de escrita e de receber direcionamentos para aprimorar sua produção acadêmica. Ele fornece insights sobre a adequação do conteúdo, a organização das ideias, a clareza da argumentação, a precisão na expressão e a adesão às normas

acadêmicas estabelecidas. Além disso, o feedback ajuda os estudantes a entenderem melhor as expectativas dos professores e avaliadores, permitindo-lhes ajustar seu desempenho e alcançar melhores resultados.

Ao receber feedback, os estudantes são incentivados a refletir sobre seu trabalho, identificar pontos fortes e áreas de melhoria, e fazer revisões e aperfeiçoamentos necessários. O feedback também contribui para a construção do conhecimento, pois incentiva os estudantes a analisar criticamente suas próprias ideias, considerar perspectivas alternativas e expandir seu repertório intelectual (Carless e Boud, 2018).

O feedback eficaz é construtivo, específico, orientado para o desenvolvimento e fornecido de forma oportuna. Ele oferece orientações claras sobre como melhorar o trabalho, apontando pontos fortes e sugerindo alternativas para aprimorar os pontos fracos. Além disso, o feedback deve ser adaptado às necessidades individuais dos estudantes, levando em consideração seu nível de proficiência, conhecimento prévio e metas de aprendizagem (Hattie e Timperley, 2007).

Uma estratégia comumente utilizada é o feedback descritivo, que se concentra em fornecer uma análise detalhada e específica sobre o trabalho do estudante. Esse tipo de feedback destaca os pontos fortes do trabalho, identifica áreas que precisam de melhorias e fornece orientações claras sobre como aprimorar o desempenho. O feedback descritivo permite que os estudantes compreendam as razões por trás das sugestões de melhoria e auxilia no processo de autorreflexão e autocorreção (Nicol e Macfarlane-Dick, 2006).

Outra estratégia é o feedback formativo, que se concentra no desenvolvimento contínuo dos estudantes, com ênfase no progresso e no aprendizado. Esse tipo de feedback é oferecido ao longo do processo de escrita, permitindo que os estudantes façam revisões e aprimorem seu trabalho com base nas orientações fornecidas. O feedback formativo encoraja os estudantes a verem os erros como oportunidades de aprendizado e a se engajarem ativamente na melhoria de suas habilidades de escrita (Higgins, Hartley e Skelton, 2002).

Além disso, a utilização de exemplos e modelos de alta qualidade pode ser uma estratégia eficaz para fornecer feedback. Ao apresentar aos estudantes trabalhos acadêmicos exemplares, eles têm a oportunidade de analisar as características de escrita bem-sucedidas, aprender com esses exemplos e aplicar essas práticas em seu próprio trabalho. Os exemplos e modelos podem servir como referência e inspiração para os estudantes, auxiliando-os na compreensão das expectativas acadêmicas e na melhoria de suas habilidades de escrita (Carless, 2007).

O professor, com sua experiência e conhecimento do conteúdo, pode fornecer uma análise mais aprofundada e contextualizada da escrita dos estudantes. Esse tipo de feedback é personalizado e adaptado às necessidades individuais de cada estudante, levando em consideração seu nível de proficiência, conhecimento prévio e objetivos de aprendizagem. O feedback do professor pode oferecer considerações valiosas sobre a organização do texto, a lógica argumentativa, a clareza da expressão e a adesão às normas acadêmicas estabelecidas. Além disso, o feedback do professor permite uma interação direta com os estudantes, abrindo espaço para esclarecimentos, discussões e orientações adicionais (Hattie e Timperley, 2007).

Carless (2015) afirma que, além de auxiliar os alunos na correção e melhoria de seus textos, o feedback avaliativo também contribui para o desenvolvimento da autoconfiança e da motivação dos alunos em relação à escrita.

Porém, uma das dificuldades enfrentadas por professores para fornecer feedbacks muito detalhados para seus alunos, consiste justamente na quantidade de alunos e trabalhos escritos que necessitam de avaliação e retorno. Algumas pesquisas têm abordado essa questão, como o estudo de Adam et al. (2019), que destaca a sobrecarga de trabalho como um dos fatores que podem prejudicar a qualidade da avaliação das atividades dos alunos.

Outra pesquisa relevante é a de Sousa et. al. (2018), que aponta a falta de tempo como um dos principais obstáculos para uma avaliação mais detalhada e criteriosa dos trabalhos. Esses autores destacam a importância de se buscar alternativas para lidar com essa dificuldade, como a utilização de ferramentas digitais de correção automática.

No campo da comparação entre o feedback automático e o feedback do professor, é possível encontrar defesas de ambos os lados, o trabalho de Taskiran e Goksel (2022) examinou como os dois tipos de feedback contribuem para o desempenho da escrita acadêmica, os resultados apresentados indicam que ambos os tipos de feedbacks foram capazes de melhorar a escrita dos alunos, tendo o feedback do professor maior efeito em alguns aspectos.

Wang e Han (2022) apresentaram em seu estudo a comparação entre os efeitos do feedback do professor e do feedback automatizado por software, como resultado, citam os efeitos positivos do feedback do professor no campo psicológico da escrita, enquanto o feedback automatizado teve maior eficácia nos aspectos cognitivos.

Já o trabalho de Wilson et. al. (2014) comparou a eficácia do feedback automático com o feedback de pares. Assim como os demais, esse trabalho reconhece o benefício de ambos os feedbacks, mas apresenta como resultados as vantagens ao feedback automático em quesitos como velocidade de devolução, disponibilidade de acesso e objetividade na resposta.

No projeto aqui proposto, buscou-se uma abordagem de “humano no processo” (do inglês “human in the loop”) no que diz respeito à inserção do professor no processo de avaliação e devolução de feedback para os alunos no contexto da escrita de projetos de pesquisa. Tal abordagem justifica-se quando consideramos as fragilidades no estabelecimento de uma comunicação restrita entre aluno-computador, com apoio da inteligência artificial, como discutido por SEO et al. (2021). Os autores apontam algumas vantagens nessa dinâmica, como a possibilidade de escalabilidade do processo, mas também apontam problemas relacionados ao risco de violação de limites sociais, quando levados em consideração aspectos de responsabilidade, autonomia e supervisão.

Com isso, preocupa-se em como os feedbacks de professores podem ser mais satisfatórios e personalizados aos alunos. Gallien e Early (2005) descobriram em seu estudo que o feedback personalizado de professores obteve nos alunos, maior satisfação e desempenho acadêmico, quando comparado com alunos que receberam apenas feedbacks automatizados. Os autores ainda vinculam a

construção de conexões mais fortes entre alunos e professores através do feedback personalizado.

2.1.4 Feedback automatizado e feedback personalizado

A revisão de trabalhos de pesquisa envolve a identificação e correção de problemas diversos, entre eles problemas relacionados à estrutura, clareza e coesão do texto, além da avaliação do conteúdo e da metodologia empregados. Para isso, é necessário que o revisor tenha um conhecimento prévio das normas técnicas e das características do gênero acadêmico em questão (Barbosa, 2017).

No contexto da escrita de projetos de pesquisa, o feedback personalizado fornecido por um professor considera o contexto específico do texto e oferece sugestões de melhoria com base em critérios específicos, como objetivos de pesquisa, metodologia e estilo acadêmico.

Segundo Barbosa (2017), o revisor pode ser visto como um colaborador experiente que auxilia o autor na construção de seu texto e no aprimoramento de sua capacidade de produção de trabalhos acadêmicos. A revisão, portanto, pode ser entendida como um processo colaborativo que envolve tanto o revisor quanto o autor na busca por uma produção mais adequada e de qualidade.

Dessa forma, o professor é quem vai apontar os erros e fazer as indicações de caminhos e correções necessárias, orientando o estudante quanto às expectativas e normas da escrita acadêmica e fornecendo feedbacks construtivos ao longo do processo de produção do trabalho.

Nesse sentido, ao orientar o aluno na escrita de seu projeto de pesquisa, o professor pode ajudá-lo a identificar lacunas em sua compreensão teórica, problemas de metodologia e outras questões que podem afetar a qualidade do trabalho final.

Barbosa (2017) complementa que o professor pode atuar como mediador entre o aluno e um possível revisor externo ou automatizado, auxiliando na interpretação das sugestões e críticas recebidas, incentivando e orientando o aluno a buscar soluções necessárias.

Por outro lado, o uso de ferramentas computacionais pode auxiliar o professor no fornecimento desses feedbacks, permitindo que pontos específicos possam ser automatizados, como a utilização de verbos nos objetivos ou a questão de pesquisa, possibilitando que o professor se concentre em pontos que o software não consegue analisar. Assim, feedbacks fornecidos por software podem ser uma ferramenta complementar para auxiliar no processo de avaliação e feedback de trabalhos de pesquisa.

A avaliação de projetos de pesquisa desempenha um papel crucial no avanço do conhecimento científico. Ela envolve uma análise crítica e sistemática do conteúdo, metodologia, resultados e contribuições de um estudo acadêmico. Através dessa avaliação, busca-se determinar a qualidade, originalidade e relevância do trabalho, além de identificar áreas de melhoria e promover o desenvolvimento contínuo da pesquisa.

A importância da avaliação de projetos de pesquisa é multifacetada. Em primeiro lugar, ela garante a validação e confiabilidade dos resultados científicos, estabelecendo padrões rigorosos e critérios de avaliação. De acordo com Lillis e Curry (2010), a avaliação é um componente fundamental para garantir a credibilidade e a validade dos resultados científicos. Através desse processo, é possível verificar se os métodos utilizados são apropriados, se os resultados são consistentes e se as conclusões estão fundamentadas em evidências sólidas. Na avaliação de projetos acadêmicos, isso não é muito diferente, mesmo que muitas vezes os projetos sejam ainda propostas não tão consistentes e não tenham propriamente resultados.

Casnici e Grimaldo (2019) ressaltam que a avaliação de projetos e relatos de pesquisa contribui para a melhoria da qualidade dos trabalhos, através da identificação de erros, inconsistências e lacunas conceituais. Ao fornecer um feedback construtivo, os avaliadores ajudam os pesquisadores a aprimorar seus estudos, fortalecendo a base de conhecimento existente.

Outro aspecto relevante da avaliação de trabalhos de pesquisa é sua influência na carreira acadêmica dos pesquisadores. De acordo com Lee et al. (2013), a avaliação de trabalhos é frequentemente utilizada como critério para promoção, concessão de financiamento e reconhecimento na comunidade científica.

Portanto, receber um feedback positivo e ter trabalhos avaliados de forma favorável pode impactar significativamente a progressão profissional dos pesquisadores.

Assim, entende-se que a avaliação e o feedback na escrita acadêmica são processos multifacetados, que requerem um equilíbrio entre as abordagens pessoais e tecnológicas. O papel do professor como mediador e orientador é crucial, mas as ferramentas tecnológicas também oferecem oportunidades significativas para melhorar e agilizar o processo de avaliação. A combinação de feedback personalizado e automatizado pode criar um ambiente de aprendizagem mais rico e eficiente, promovendo o desenvolvimento de habilidades de escrita acadêmica de alta qualidade e o avanço do conhecimento científico.

2.2 A ANALÍTICA DA APRENDIZAGEM

A analítica da aprendizagem, ou *Learning Analytics* em inglês, envolve a coleta e análise de dados produzidos por alunos no processo de aprendizagem, buscando identificar modos e estratégias mais eficazes e como o professor pode melhor apoiar os alunos em suas jornadas educacionais. Assim, esse campo explora padrões e tendências dentro dos dados educacionais para melhor compreender e apoiar o aprendizado do aluno, promovendo um ensino mais adaptativo e centrado no aluno (Siemens e Baker, 2012).

Uma das valiosas contribuições da Analítica da Aprendizagem é sua capacidade de fornecer ideias e visualização de padrões que podem levar a uma personalização mais efetiva da educação. Através da análise de dados, professores, gestores, pais e os próprios alunos podem identificar os pontos fracos e fortes na aprendizagem, adaptando suas estratégias de ensino para atender às necessidades individuais (Dawson, Gašević, Siemens, e Joksimović, 2018). Essa abordagem oferece a promessa de melhorar a eficácia do ensino, aumentar o engajamento do aluno e, em última análise, melhorar os resultados educacionais.

Cabe ressaltar que com o aumento expressivo do número de estudantes na modalidade EaD, ocorre uma diminuição do contato aluno-professor mas, em contrapartida, aumenta-se a produção de dados educacionais por meio de Ambientes Virtuais de Aprendizagem (AVAs) (Almosallam e Ouertani, 2014). Assim,

considerando que os dados já estão produzidos e disponibilizados, não utilizá-los de maneira adequada, ou seja, visando o aprimoramento do processo de aprendizagem é um desperdício de oportunidades de buscar melhorar a educação.

Corroborando com isso, Siemens e Baker (2012) colocam que os dados produzidos em AVAs representam uma importante fonte de informação, capazes de informar práticas pedagógicas com vistas ao aprimoramento do processo de aprendizagem, permitindo uma compreensão detalhada das atividades dos estudantes.

Assim, compreende-se a utilização da Analítica da Aprendizagem como uma possibilidade valiosa no fornecimento de informações aos educadores sobre as necessidades individuais dos alunos, permitindo que adaptem suas abordagens pedagógicas de maneira mais eficaz (Long e Siemens, 2011).

Nesse contexto é que se concentra o campo da Analítica da Aprendizagem, fazendo uso de técnicas de mineração e análise de dados, estatísticas e sistemas de informação, os quais buscam obter insights sobre a aprendizagem dos alunos em diferentes ambientes e contextos, auxiliando os envolvidos numa compreensão aprimorada do processo de aprendizagem por meio de técnicas que coletam, processam e fornecem visualização de dados.

O trabalho de Campos et. al. (2020) apresenta uma verificação das contribuições e usos da Analítica da Aprendizagem no contexto educacional brasileiro. Os autores efetuaram buscas em três revistas na área de Informática na Educação e em anais de dois eventos de relevância nacional. Foram encontrados 136 artigos entre os anos de 2018 e 2020. Como resultado, o trabalho apresenta que o desempenho acadêmico dos alunos foi abordado em 36,62% dos artigos selecionados, seguido pelo problema da evasão escolar (21,13%), aplicação de técnicas para classificação de perfis de alunos (8,45%), aspectos de motivação e engajamento de alunos (7,04%), e interações de tutores em ambientes virtuais de aprendizagem (4,23%). Para finalizar as informações apresentadas, é informado que a maioria dos autores direcionou suas pesquisas para a modalidade de Educação a Distância (36,62%), o que os autores atribuem à facilidade de coleta de dados, já que essa modalidade exige o uso de AVAs.

As dificuldades e deficiências em leitura e escrita por parte dos alunos brasileiros são expostas através dos resultados de diversos exames e avaliações divulgados temporalmente. No ensino básico e fundamental, avaliações como o ANA mostram que cerca de 34% dos alunos se encontram em níveis insuficientes de escrita, sendo que apenas 8,28% dos alunos avaliados atingiram o índice “Desejável” nessa competência. Já segundo o MEC, no ENEM de 2021, a média de pontuação na redação foi de 634,16 pontos, tendo 22 candidatos obtido nota máxima e 95.788 candidatos nota zero, sendo que 2.267.350 candidatos compareceram à prova. Com esses problemas em foco, o processo de escrita torna-se alvo de pesquisas acadêmicas variadas. Como início, é necessário identificar e compreender os conceitos da analítica da aprendizagem e como essa pode auxiliar o professor a melhor abordar as dificuldades de seus alunos.

2.2.1 Analítica da escrita

A Analítica da Escrita, subárea da Analítica da Aprendizagem, também concentra-se na coleta e análise de dados relacionados a aprendizagem dos alunos, porém, é mais restrita quanto ao tipo, focando especificamente no acompanhamento do processo de produção textual, buscando possibilitar uma visão melhorada do progresso da habilidade de escrita dos alunos.

Essa área especializada aplica métodos analíticos para avaliar e aprimorar as habilidades de escrita, utilizando avanços em tecnologias de processamento de linguagem natural e análise de texto. Segundo Shermis e Burstein (2013), tais tecnologias permitem uma avaliação mais profunda e objetiva dos textos dos alunos, fornecendo feedback preciso sobre diversos aspectos da escrita.

Assim, essa tecnologia aumenta a capacidade do profissional da educação em monitorar o desenvolvimento da escrita de seus alunos. A Analítica da Escrita fornece aos educadores, ferramentas valiosas para intervenções pedagógicas personalizadas. Crossley, McNamara e McCarthy (2010), por exemplo, afirmam que a análise de estrutura, coerência, vocabulário e gramática pode revelar deficiências e padrões na aprendizagem da escrita.

Além disso, a Analítica da Escrita pode beneficiar os alunos ao fornecer feedback imediato e detalhado sobre seus textos. Essa capacidade de análise automatizada, como discutida por Warschauer e Ware (2006), oferece aos alunos a oportunidade de receberem uma avaliação rápida de seus trabalhos, promovendo a autoavaliação e o aprimoramento contínuo das habilidades de escrita.

Porém, aplicar a Analítica da Escrita no contexto educacional expõe certos cuidados necessários e requisitos críticos. Primeiramente, a precisão e a confiabilidade das ferramentas de análise de texto são fundamentais. Como indicado por Valenti, Neri e Cucchiarelli (2003), a precisão na avaliação de textos escritos por alunos é crucial para garantir que o feedback fornecido seja válido e útil. As ferramentas de analítica devem ser capazes de identificar corretamente não apenas erros gramaticais e ortográficos, mas também avaliar aspectos mais complexos, como coesão, coerência e argumentação.

Outro requisito importante é a adaptabilidade das ferramentas de Analítica da Escrita a diferentes níveis de habilidade de escrita e a variados estilos de texto. Como Sherri e Burstein (2013) observam, diferentes alunos possuem diferentes níveis de proficiência em escrita, e as ferramentas devem ser sensíveis a essas variações. A capacidade de se adaptar a diferentes estilos de escrita, sejam eles acadêmicos, criativos ou técnicos, é fundamental para que a analítica seja eficaz em diversos contextos educacionais.

Além disso, a integração das ferramentas de Analítica da Escrita no ambiente educacional deve ser feita de forma complementar e não como substituto do papel do educador. Como argumentado por Warschauer e Ware (2006), enquanto as ferramentas de analítica oferecem uma avaliação objetiva e imediata, a interpretação e a intervenção pedagógica de um educador são insubstituíveis. O educador deve ser capaz de utilizar as informações fornecidas pela analítica para orientar e personalizar o ensino.

A usabilidade das ferramentas de Analítica da Escrita também é um fator crítico. Como apontado por Burstein, Chodorow e Leacock (2004), as ferramentas devem ser acessíveis e fáceis de usar tanto para educadores quanto para alunos. Uma interface de usuário intuitiva e um processo de feedback claro são essenciais

para garantir que tanto os alunos quanto os professores possam efetivamente utilizar as ferramentas em seu benefício.

Por fim, a questão da privacidade e segurança dos dados dos alunos é de suma importância. Conforme destacado por Cope e Kalantzis (2016), ao lidar com dados de escrita dos alunos, é crucial garantir que as informações sejam manuseadas e armazenadas com o máximo de segurança e privacidade. As escolas e instituições devem adotar políticas claras e rigorosas de proteção de dados para manter a confiança dos alunos e dos pais no uso dessas tecnologias.

2.2.2 Processamento de Linguagem Natural (PLN)

O Processamento de Linguagem Natural (PLN) é um campo interdisciplinar que combina elementos da linguística computacional, inteligência artificial e ciência da computação para permitir que máquinas compreendam, interpretem e manipulem a linguagem humana (Jurafsky e Martin, 2018).

O PLN teve início com a criação de modelos linguísticos simples, evoluindo para sistemas complexos de aprendizado de máquina. Nos primórdios, o PLN baseava-se em regras e abordagens simbólicas (Hutchins, 2006), progredindo para modelos estatísticos na década de 1980 e culminando nos recentes avanços em aprendizado profundo e redes neurais (Bengio, Courville e Vincent, 2013). Esse progresso não apenas ampliou a capacidade de processamento linguístico das máquinas, mas também abriu novos caminhos para aplicações práticas.

Hoje, o PLN encontra-se em inúmeras aplicações, transformando a maneira como interagimos com a tecnologia. Desde os sistemas de recomendação personalizados até os sofisticados chatbots, o PLN está no cerne de muitas inovações tecnológicas (Hovy e Lavid, 2010). Essas aplicações não apenas demonstram o potencial do PLN em diversos setores, mas também ressaltam a sua crescente importância no mundo tecnologicamente interconectado de hoje.

No campo educacional, dentre as diversas aplicações, destacam-se as tentativas de previsão de níveis e notas baseadas em métricas textuais. Trabalhos como os de Xia et. al. (2016) e Pires et. al. (2017) mostram aplicação da tecnologia

para a avaliação do nível de leitura e complexidade dos textos, objetivando direcioná-los para a etapa escolar e/ou nível de compreensão apropriado.

Outro uso educacional recorrente na literatura é a identificação ou simplificação de conteúdos textuais para diferentes anos escolares, Crossley et. al. (2010) e Aluísio et. al. (2010) apontam que esse uso pode auxiliar na seleção ou transformação de material em conteúdos mais acessíveis e envolventes para as faixas etárias e os diversos níveis de desenvolvimento da leitura.

Evers e Finatto (2016) aplicaram técnicas de aprendizado de máquina e processamento de linguagem natural para identificar diferentes níveis de proficiência em português brasileiro. Elas se concentraram em padrões lexicais e de coesão, e seu estudo revelou que, ao utilizar o Coh-Metrix-Port, ferramenta para calcular métricas textuais, um conjunto específico de métricas apresentou padrões consistentes e significativos, correspondentes às categorias de proficiência estabelecidas.

Martins (2016) conduziu uma pesquisa examinando a relação entre medidas de complexidade textual e o sucesso acadêmico dos estudantes. Os achados do estudo apontaram para uma correlação positiva entre a frequência de uso de substantivos e o progresso acadêmico, enquanto uma correlação negativa foi observada entre a frequência de verbos e a progressão. Quanto à complexidade sintática, a pesquisa descobriu que a correlação variava conforme o tipo de oração subordinada, diferindo nos dois conjuntos de dados examinados.

Esses estudos apresentam e corroboram as capacidades da Analítica da Escrita, bem como das tecnologias de PLN, na aplicação em contextos educacionais, auxiliando alunos e professores na identificação de pontos fortes e fracos, possibilitando, assim, melhora nas habilidades de leitura e escrita.

2.2.2.1 Ferramentas de PLN

As aplicações do PLN crescem juntamente com o avanço tecnológico, um exemplo atual é a explosão de modelos de larga escala (Large Language Models - LLMs), como GPT (Generative Pre-trained Transformer) da OpenAI. Esses modelos,

através do treinamento em grandes conjuntos de dados, passaram a possuir uma capacidade notável de gerar texto, responder perguntas e até mesmo criar conteúdo em diversos formatos (Brown et al., 2020). A utilidade dos LLMs tem sido demonstrada em uma variedade de aplicações, desde a geração automatizada de texto até a assistência em tarefas de programação (Radford et al., 2019).

Outra ferramenta crucial em PLN é a análise de sentimentos. Essa técnica utiliza algoritmos de aprendizado de máquina para identificar e categorizar opiniões expressas em textos, ajudando a entender as reações emocionais dos usuários (Liu, 2012). Através dessa análise, é possível monitorar e responder a feedback de clientes em mídias sociais e revisões de produtos, proporcionando insights valiosos para estratégias de marketing e desenvolvimento de produtos (Pang e Lee, 2008).

O Reconhecimento de Entidades Nomeadas (NER) é outra ferramenta importante, permitindo a identificação automática de entidades como nomes de pessoas, organizações, locais e expressões temporais em textos. O NER é amplamente utilizado em sistemas de extração de informações, melhorando a capacidade de recuperação e organização de dados em grandes conjuntos de textos (Nadeau e Sekine, 2007). Essa tecnologia tem aplicações significativas em áreas como a extração de informações biomédicas e análise de mídia.

A tradução automática, potencializada por modelos de PLN como o Transformer, demonstrou avanços significativos, reduzindo a barreira linguística em comunicações globais (Vaswani et al., 2017). Modelos multilíngues, como o mBERT e o XLM-R, oferecem uma compreensão e geração de linguagem em múltiplos idiomas, facilitando a comunicação e o acesso a informações em diversas línguas (Devlin et al., 2018; Conneau et al., 2020).

Focando-se no escopo deste trabalho, ferramentas de PLN podem ser utilizadas no processamento de trabalhos acadêmicos de diversas formas e para diversas finalidades, abrangendo desde ferramentas de detecção de plágio à sumarização e avaliação da qualidade e complexidade da escrita.

No âmbito das ferramentas de detecção de plágio, o PLN dá a esses sistemas a capacidade de identificar não apenas cópias literais, mas também parafraseamentos e plágios mais sutis. Rosales et. al. (2008) apresentam em seu

trabalho uma ferramenta computacional com esse propósito, descrevendo as necessidades como um banco de trabalhos para comparação de textos e os passos dessa comparação.

Outro uso importante do PLN é na avaliação da complexidade da escrita e da coerência em textos acadêmicos. Ferramentas de PLN podem analisar a estrutura gramatical, a clareza, a fluidez e a coerência lógica dos textos, fornecendo feedbacks valiosos para os autores. Reid (2016) diz que tais sistemas podem utilizar técnicas de análise sintática e semântica para avaliar a complexidade e a clareza da escrita, argumentando que essa análise pode ser útil para autores não nativos em inglês.

Porém, para ser possível avaliar a complexidade de um texto, se faz necessário capacitar o software a reconhecer, identificar e padronizar trechos de textos e palavras, buscando assim seu significado e pertencimento dentro do contexto. Essa ação de capacitação exige o uso de ferramentas de PLN para reduzir as palavras à sua forma mais básica e segmentar um texto em unidades menores.

2.2.2.2 Tokenizadores

No universo do PLN, diversas ferramentas e tratamentos se fazem necessários. Os tokenizadores surgem como ferramentas essenciais, atuando como a ponte entre o texto bruto e sua transformação em dados estruturados e analisáveis por máquinas. Após o tópico anterior ter abordado as ferramentas gerais de PLN, é fundamental entender como os tokenizadores operam dentro desse ecossistema, desempenhando um papel crítico na preparação inicial dos dados. Eles são responsáveis por decompor o texto em elementos menores, conhecidos como tokens, que podem variar de palavras individuais a caracteres ou subunidades linguísticas. Essa segmentação inicial é crucial para qualquer aplicação de PLN, pois estabelece a base para análises mais sofisticadas, como reconhecimento de entidades nomeadas, análise de sentimentos e tradução automática. A eficiência e precisão de um tokenizador, portanto, são determinantes para o sucesso subsequente de todas as operações de PLN.

A eficácia de um algoritmo de tokenização depende significativamente da complexidade da língua e do contexto do texto. Por exemplo, em línguas como o inglês, a tokenização pode ser relativamente direta, focando na separação de palavras com base em espaços e pontuação. Contudo, em línguas como o chinês, onde não há espaços claros entre as palavras, a tokenização requer métodos mais sofisticados para identificar corretamente os limites das palavras (Devlin et. al., 2018). Além disso, a tokenização também deve lidar com desafios como a identificação de entidades nomeadas, abreviações e construções gramaticais complexas.

Os avanços em algoritmos de tokenização têm desempenhado um papel crucial no aprimoramento de modelos de PLN. Modelos modernos, como o BERT (Bidirectional Encoder Representations from Transformers) e o GPT (Generative Pre-trained Transformer), utilizam técnicas avançadas de tokenização para melhor capturar as nuances linguísticas. Ainda segundo Devlin et. al. (2018), o BERT utiliza uma abordagem chamada WordPiece, que divide palavras em subpalavras, permitindo que o modelo lide eficientemente com um vocabulário amplo sem a necessidade de armazenar todas as palavras possíveis.

Uma ferramenta de tokenização reconhecida e utilizada em projetos como o NILC-Matrix (Leal, 2023) é o Natural Language Toolkit (NLTK). Essa ferramenta é uma biblioteca de referência em linguagem Python, sendo amplamente utilizada no campo do Processamento de Linguagem Natural (PLN). Entre suas muitas funcionalidades, os tokenizadores do NLTK são particularmente notáveis pela sua versatilidade e eficácia em lidar com diferentes tipos de textos e línguas.

O NLTK oferece uma variedade de tokenizadores, cada um projetado para atender a diferentes necessidades e tipos de texto. Por exemplo, o `word_tokenize` é um dos mais utilizados e é baseado no tokenizador Penn Treebank, eficaz na separação de palavras e pontuação em idiomas como o inglês (Bird, Loper e Klein, 2009). Para textos que requerem uma abordagem mais granular, como a análise de tweets, o NLTK fornece o `TweetTokenizer`, que é mais eficaz na manipulação das peculiaridades desse tipo de texto, como hashtags e menções a usuários.

Um aspecto promissor do NLTK em relação ao português brasileiro é sua capacidade de integração com outras bibliotecas e recursos de PLN disponíveis em

linguagem Python. Por exemplo, pode-se combinar o NLTK com a biblioteca spaCy para acessar modelos de processamento de linguagem específicos para o português. Essa integração permite que pesquisadores e desenvolvedores aproveitem os pontos fortes de ambas as bibliotecas, melhorando a precisão e eficiência do processamento de textos em português brasileiro (Jugran et. al. 2021).

A importância do NLTK no campo do PLN é amplamente reconhecida, tanto na academia quanto na indústria. Sua flexibilidade, a extensa documentação e a comunidade ativa contribuem para seu status como uma ferramenta indispensável para pesquisadores e desenvolvedores de PLN. O NLTK não só facilita a execução de tarefas complexas de PLN, mas também serve como um recurso educacional valioso para aqueles que estão aprendendo os fundamentos do campo (Bird, Klein, e Loper, 2009).

2.2.2.3 Parsers

Após a análise dos tokenizadores no Processamento de Linguagem Natural (PLN), a atenção se volta para a função dos parsers, uma ferramenta fundamental na compreensão e análise estrutural de textos. Enquanto os tokenizadores dividem o texto em unidades menores, os parsers assumem a responsabilidade de analisar a estrutura gramatical e semântica desses tokens, desempenhando um papel crucial na interpretação e entendimento do texto.

A função primordial de um parser em PLN é decompor sentenças em sua estrutura gramatical e identificar as relações entre as palavras. Essa análise é crucial para entender a organização e o significado das sentenças. Um parser bem-sucedido transforma uma sequência linear de tokens em uma estrutura hierárquica que reflete a organização gramatical do texto (Jurafsky e Martin, 2018). Essa estrutura é muitas vezes representada na forma de uma árvore de análise, onde os nós representam as palavras e as conexões entre eles representam relações gramaticais, como sujeito, predicado e objeto.

No campo do PLN, existem dois tipos principais de parsers: os sintáticos e os semânticos. Os parsers sintáticos, como o Stanford Parser (Klein e Manning, 2003), concentram-se em analisar a estrutura gramatical das sentenças. Eles são

fundamentais para aplicações que necessitam entender a estrutura gramatical de um texto, como na análise de sentimentos ou na extração de informação. Por outro lado, os parsers semânticos, como o SEMAFOR (Das et al., 2014), visam a compreender o significado das sentenças, analisando as relações semânticas entre as palavras, o que é vital em sistemas de perguntas e respostas e em tradução automática.

A evolução dos parsers tem sido uma das áreas mais dinâmicas no PLN. Com o advento das Redes Neurais Profundas e do aprendizado de máquina, surgiram parsers mais avançados e precisos. Por exemplo, o uso de Transformers, como no modelo BERT (Devlin et al., 2018), revolucionou o parsing semântico ao fornecer uma compreensão contextual mais profunda das sentenças.

Avançando na discussão sobre parsers no Processamento de Linguagem Natural (PLN), é fundamental destacar as especificidades e desafios enfrentados ao aplicar essas ferramentas à língua portuguesa. Diferentemente do inglês, que tem recebido ampla atenção no desenvolvimento de parsers, o português apresenta desafios únicos que exigem soluções adaptadas e específicas.

Um dos principais desafios no parsing do português é sua rica morfologia e sintaxe flexível, que podem complicar a análise gramatical e semântica. Por exemplo, a ordem das palavras em uma sentença em português pode variar mais do que em inglês, sem alterar o significado. Isso requer parsers que sejam capazes de entender e interpretar essa flexibilidade (Aluísio et al., 2003). Além disso, a concordância de gênero e número, comum em português, adiciona uma camada extra de complexidade na análise morfológica e sintática.

Para abordar essas questões, foram desenvolvidos parsers específicos para o português, como o PALAVRAS (Bick, 2000), um parser morfossintático amplamente utilizado para o português europeu e brasileiro. O PALAVRAS oferece análise detalhada da estrutura gramatical de textos em português, incluindo a identificação de classes gramaticais, funções sintáticas e estruturas de frases. Esse parser tem sido uma ferramenta valiosa para pesquisadores e profissionais trabalhando com PLN em português, fornecendo insights detalhados sobre a complexidade da língua.

Além disso, projetos de colaboração internacional, como o Universal Dependencies (Nivre et al., 2016), têm contribuído para o avanço do parsing em línguas como o português. O Universal Dependencies fornece um framework e um conjunto de dados anotados para várias línguas, incluindo o português, facilitando o desenvolvimento de parsers que são consistentes e comparáveis entre diferentes idiomas. Isso permite uma análise mais uniforme e abrangente, fundamental para o avanço do PLN em um contexto multilíngue.

Em resumo, os parsers são instrumentos indispensáveis no PLN, permitindo que máquinas não apenas processem, mas também compreendam a estrutura e o significado do texto humano. Sua aplicação vai desde a análise básica de textos até tarefas complexas de compreensão de linguagem, demonstrando sua relevância e versatilidade no campo.

2.2.2.4 Lematizadores

Após a discussão sobre tokenizadores e parsers no contexto do PLN, é importante abordar a função dos lematizadores, que desempenham um papel essencial na normalização e interpretação linguística dos dados. Lematizadores são sistemas projetados para reduzir palavras à sua forma de lema, isto é, a forma base ou de dicionário, considerando a função gramatical e o contexto em que aparecem (Manning e Schütze, 1999). Essa etapa é fundamental no cronograma de PLN, pois facilita o tratamento de variações morfológicas e aumenta a eficiência de processos subsequentes, como análise semântica e classificação de texto.

A função dos lematizadores contrasta com a dos simples removedores de sufixos (stemmers), pois os lematizadores incorporam uma compreensão mais profunda da morfologia da língua. Por exemplo, em inglês, as palavras "running", "ran", e "runs" seriam lematizadas para "run", enquanto um removedor de sufixos poderia simplesmente reduzi-las a "runn" (Jurafsky e Martin, 2018). Essa distinção é crucial para o entendimento preciso do texto, especialmente em idiomas com rica variação morfológica.

Diversas ferramentas de lematização têm sido desenvolvidas para facilitar esse processo em vários idiomas. Por exemplo, a biblioteca spaCy, já citada

anteriormente, que fornece, além de tokenizadores, lematizadores para múltiplas línguas, utilizando modelos de linguagem que levam em conta o contexto lexical e sintático (Jugran et. al., 2021). Essas ferramentas demonstram a aplicabilidade prática da lematização em diversos cenários de PLN, desde a análise de sentimentos até o processamento de linguagem em chatbots.

Em suma, tokenizadores, parsers e lematizadores constituem a espinha dorsal do PLN. Eles permitem que máquinas processem e compreendam a linguagem humana de maneira cada vez mais sofisticada, abrindo novas possibilidades para a interação humano-computador e o processamento automatizado de grandes volumes de texto. A exploração contínua e aprimoramento dessas ferramentas são fundamentais para o avanço contínuo do campo do PLN e suas inúmeras aplicações no mundo real.

2.3 A AVALIAÇÃO DE PROJETOS DE PESQUISA

A avaliação de projetos de pesquisa é um aspecto crucial no avanço do conhecimento científico e tecnológico. Essa avaliação desempenha um papel fundamental em diversas fases do ciclo de vida de um projeto, desde a concepção inicial até a análise de resultados e impacto. O primeiro passo na avaliação de projetos de pesquisa envolve a definição clara de objetivos e a verificação de sua relevância. Kothari (2004) enfatiza a importância de objetivos bem definidos e alinhados com as lacunas de conhecimento existentes.

A metodologia é outro pilar na avaliação de projetos de pesquisa. A escolha da metodologia adequada é crucial para garantir a validade e confiabilidade dos resultados. Yin (2014) destaca a importância de selecionar métodos de pesquisa que se alinhem aos objetivos do estudo e às questões de pesquisa. A adequação da metodologia influencia diretamente a interpretação e a utilidade dos resultados obtidos.

Além disso, a competência e a experiência da equipe de pesquisa são essenciais para a execução bem-sucedida de um projeto. Booth et. al. (2016) discutem a importância de uma equipe com habilidades complementares e experiência adequada para enfrentar os desafios que surgem durante a pesquisa.

Aspectos éticos e de conformidade também são indispensáveis em qualquer projeto de pesquisa. Resnik (1998) discute a importância de se aderir a princípios éticos na condução de pesquisas, incluindo o respeito por normas regulatórias e a integridade dos dados.

2.3.1 Critérios de avaliação de projetos de pesquisa

A seleção dos critérios de avaliação adequados desempenha um papel fundamental na análise objetiva e precisa dos projetos de pesquisa. A definição desses critérios deve considerar a relevância científica, a originalidade, a consistência metodológica e o potencial de contribuição do projeto para a área de estudo.

Ao selecionar os critérios de avaliação, é importante considerar as especificidades do campo de pesquisa, bem como as diretrizes e padrões estabelecidos pelas instituições ou agências financiadoras. A definição adequada dos critérios proporciona uma estrutura sólida para a avaliação e permite uma análise mais objetiva e sistemática do projeto.

De acordo com Tinkler e Jackson (2004), a originalidade e a relevância da proposta são critérios-chave na avaliação de projetos de pesquisa. A originalidade garante a contribuição para o conhecimento existente, enquanto a relevância assegura que o projeto aborde questões pertinentes e significativas dentro da área de estudo.

A coerência teórica também desempenha um papel importante na avaliação. Conforme destacado por Creswell (2014), a base teórica deve fornecer um arcabouço conceitual sólido para sustentar o projeto de pesquisa, demonstrando a compreensão dos fundamentos teóricos e sua aplicação adequada.

A formulação clara da questão de pesquisa e dos objetivos é outro critério crucial. Segundo Rudestam e Newton (2015), a questão de pesquisa deve ser claramente definida, específica e orientada para a investigação do problema em questão. Os objetivos do projeto devem estar alinhados com a questão de pesquisa e indicar os resultados esperados.

Para avaliar o item objetivo de um projeto, Bastiani (2022) sugere a utilização de critérios de avaliação baseados na literatura sobre metodologia da pesquisa e escrita de projetos de pesquisa. Dentre os aspectos a serem observados, destacam-se:

- Existência de verbos na forma infinitiva, e se ele se restringe a essa ação, de forma a manter a precisão do objetivo em apenas uma linha de investigação;
- Se o elemento objetivo do projeto menciona, ou não, aspectos metodológicos da pesquisa, o que normalmente não é necessário nesta parte do texto;
- A presença de pronomes indefinidos no objetivo, ou termos que por definição, se referem a algo ou alguém de maneira indeterminada e imprecisa.

Segundo Creswell (2014), a formulação de uma questão de pesquisa envolve a identificação de variáveis-chave e a delimitação do escopo do estudo. O autor enfatiza que a questão de pesquisa deve ser clara, concisa e específica, evitando ambiguidades e generalizações excessivas.

Swales e Feak (2012) destacam a importância de formular questões de pesquisa orientadas para a solução de problemas. Eles sugerem que as questões de pesquisa devem abordar lacunas na literatura existente, contribuindo para o avanço do conhecimento e fornecendo respostas práticas para problemas ou desafios.

Além disso, Gil (2017) ressalta a necessidade de evitar questões de pesquisa que sejam muito amplas ou muito restritas. O autor recomenda que as questões sejam adequadas ao tempo e aos recursos disponíveis, considerando a viabilidade de pesquisa.

Outra sugestão apresentada em Bastiani (2022), para a avaliação do item questão de pesquisa, destaca os seguintes pontos:

- Se existe um, e somente um, termo interrogativo;
- Se o item está formatado como uma questão, respeitando a pontuação adequada;
- Se possui termos vagos e amplos, que poderiam dispersar o seu foco;

Bastiani (2022) ainda salienta a necessidade de verificação final entre a questão de pesquisa e o objetivo, visando manter as relações de coerência entre eles e o título do projeto.

A metodologia é um aspecto fundamental a ser considerado na avaliação. Conforme abordado por Bryman (2016), a metodologia deve ser adequada para responder à questão de pesquisa e deve ser descrita de forma clara e detalhada. Os métodos de coleta e análise de dados devem ser apropriados e transparentes.

Creswell (2014) enfatiza a importância de fornecer uma descrição detalhada dos métodos utilizados, de modo que outros pesquisadores possam replicar o estudo. O autor destaca que a seção da metodologia deve incluir informações sobre a abordagem de pesquisa (qualitativa, quantitativa ou mista), os participantes ou amostra, os instrumentos de coleta de dados e os procedimentos de análise.

Silverman (2016) ressalta a necessidade de justificar as escolhas metodológicas. O pesquisador deve explicar por que determinados métodos foram selecionados e como eles são apropriados para responder à pergunta de pesquisa. Essa justificativa ajuda a fortalecer a credibilidade do estudo e demonstra a compreensão das limitações e possíveis vieses metodológicos.

Ademais, Booth et al. (2016) enfatizam a necessidade de transparência na descrição dos procedimentos metodológicos. Os pesquisadores devem fornecer detalhes suficientes para que outros possam compreender e replicar o estudo, incluindo informações sobre o contexto do estudo, as etapas de coleta e análise de dados e os critérios utilizados na seleção da amostra.

Para o item metodologia, segundo Bastiani (2022), a literatura destaca oito métricas de avaliação, sendo elas:

- As duas primeiras investigam a classificação da pesquisa com relação ao objetivo, se é descritivo, experimental ou exploratório e a abordagem que será utilizada, sendo qualitativa, quantitativa ou mista.
- A terceira e quarta métricas abordam os termos apresentados, se esses caracterizam a estratégia ou modalidade de pesquisa a ser utilizada e a identificação dos procedimentos de coleta de dados.

- A quinta métrica é responsável por avaliar se a metodologia aborda quem serão os sujeitos da pesquisa
- A sexta métrica busca analisar se a metodologia descreve o período de realização do(s) estudo(s).
- A sétima e oitava métrica trabalham com a coerência entre os conceitos-chave, elemento textual e referencial teórico do projeto.

Por fim, a viabilidade do projeto também deve ser avaliada. Segundo Churchill et al. (2018), a viabilidade envolve a consideração dos recursos disponíveis, do cronograma proposto e da capacidade do pesquisador para realizar o projeto dentro das restrições e limitações.

2.3.2 Avaliação automática de projetos de pesquisa

A avaliação de projetos de pesquisa desempenha um papel fundamental na garantia da qualidade e do rigor científico das produções acadêmicas. Tradicionalmente, esse processo pode envolver revisões por pares e análises manuais, já no contexto de uma disciplina de metodologia científica ou da pesquisa, essa tarefa recai unicamente sobre o professor, que tem a tarefa de orientar seus alunos no processo da aprendizagem da escrita acadêmica. No entanto, fazer uso de tecnologias para análise de escrita pode proporcionar oportunidades e auxílio a esses professores, já que com a utilização de algoritmos e sistemas inteligentes capazes de avaliar projetos de pesquisa de forma ágil e precisa o professor poderá concentrar-se nas particularidades de cada aluno.

A área da avaliação automática está em constante evolução, o surgimento e evolução de sistemas de Inteligência Artificial e PLN possibilitam cada vez mais que ferramentas com a finalidade de avaliação automatizada tenham maior precisão e sejam empregadas em contextos cada vez mais variados. Conforme Zhang (2021), a avaliação automatizada tem diversas possibilidades de uso no contexto educacional, indo desde a avaliação de testes padronizados, como o *Test of English as a Foreign Language* e o *Graduate Record Exam*, até avaliações de redações de sala de aula.

As capacidades de PLN, conforme descritas por Jurafsky e Martin (2018), incluem a análise de relevância temática, consistência interna e originalidade dos projetos. Tais sistemas utilizam técnicas avançadas de aprendizado de máquina e processamento de linguagem para identificar padrões, tendências e anomalias nos textos submetidos. A implementação dessas tecnologias na avaliação de projetos de pesquisa não só acelera o processo, mas também introduz um nível de objetividade difícil de alcançar em avaliações puramente humanas.

No entanto, há desafios inerentes ao uso do PLN nessa esfera, como destacado por Bird et. al. (2009). Por exemplo, a compreensão de jargões específicos e a interpretação correta do contexto e da nuance em textos acadêmicos são aspectos que desafiam os sistemas de PLN. É crucial que tais sistemas sejam treinados com um conjunto diversificado de dados acadêmicos para melhorar sua capacidade de avaliação precisa e relevante.

Além disso, como apontam Bender e Friedman (2018) a questão da imparcialidade e da representatividade dos dados usados para treinar esses sistemas é de suma importância. A garantia de que as ferramentas de PLN não perpetuem vieses existentes nos dados é fundamental para a avaliação justa e equitativa de projetos de pesquisa.

Desta forma, a integração de ferramentas de PLN na avaliação de projetos de pesquisa oferece possibilidades promissoras, mas também exige consideração cuidadosa de suas limitações e desafios éticos. Como sugere Bird et al. (2009), a colaboração contínua entre especialistas em PLN e profissionais da educação é vital para o desenvolvimento de sistemas que não apenas automatizem, mas também enriqueçam o processo de avaliação de projetos de pesquisa.

Uma evolução para as ferramentas de avaliação automatizada é o fornecimento de mais do que apenas uma nota, como a apresentação de feedbacks, auxiliando assim o usuário a compreender os erros apontados. Nesse sentido, sistemas de PLN podem ser treinados para reconhecer padrões em escritas acadêmicas eficazes, auxiliando alunos na estruturação de argumentos e na organização de ideias. Isso é particularmente útil para estudantes em estágios iniciais de pesquisa, onde a estruturação de ideias é fundamental.

Ferramentas de avaliação com fornecimento de feedbacks também podem auxiliar professores, avaliadores, revisores e demais envolvidos no âmbito de projetos de pesquisa. Porém, como apresentado em Hyland e Hyland (2006), essas ferramentas devem ser utilizadas como complementares, e não em substituição à interação aluno-professor.

Assim, conclui-se que o uso de ferramentas de avaliação automatizada tende a crescer, porém sua escolha e uso devem sempre atentar para a participação e validação por professores, para que as vantagens de cada parte do processo sejam potencializadas, tornando o feedback ágil, preciso e personalizado às particularidades do aluno.

2.4 TECNOLOGIAS PARA ANÁLISE DA ESCRITA

A tecnologia tem desempenhado um papel cada vez mais importante na análise da escrita, permitindo que pesquisadores, professores e escritores avaliem a qualidade e a complexidade da escrita com mais precisão, eficiência e velocidade.

Uma forma de uso da tecnologia para análise e avaliação da escrita consiste no uso de softwares de feedback automatizados, capazes de auxiliar o escritor ou avaliador sobre a qualidade e a complexidade da escrita.

Outras formas pelas quais a tecnologia tem sido usada na análise da escrita é através da utilização de softwares de análise de texto, como o NILC-Metrix e softwares de mineração de texto, como o Sobek. Esses softwares utilizam uma variedade de recursos linguísticos e psicolinguísticos para avaliar a complexidade e a compreensibilidade do texto, gerando uma série de medidas que podem ser usadas para avaliar a complexidade da escrita.

2.4.1 Coh-Metrix e NILC-Metrix

O Coh-Metrix é um software de análise de texto com uso de PLN, que tem sido amplamente utilizado em pesquisas para avaliar a complexidade da escrita. Ele utiliza uma variedade de recursos linguísticos e psicolinguísticos para analisar a

complexidade e a compreensibilidade do texto, gerando uma série de medidas que podem ser usadas para avaliar a escrita (Graesser et al. 2004).

Assim, o Coh-Metrix avalia o texto em várias dimensões, incluindo coesão, estruturação do discurso, semântica, sintaxe e outras características linguísticas. Utilizam-se técnicas de PLN para extrair informações do texto e, em seguida, gera-se uma variedade de medidas para avaliar a complexidade do texto em cada dimensão.

Segundo o estudo de Crossley, et. al. (2011), o Coh-Metrix pode ser usado para avaliar a complexidade lexical e sintática de um texto, além de medir sua coerência e coesão. O software também pode ser usado para identificar palavras-chave e temas em um texto, bem como para avaliar a legibilidade e a compreensibilidade do texto.

Outro estudo, realizado por Graesser et. al. (2004), utilizou o Coh-Metrix para analisar a complexidade de textos escritos por estudantes. Os resultados mostraram que as medidas geradas pelo software foram úteis para avaliar a complexidade da escrita dos estudantes, bem como para fornecer métricas aos professores.

O trabalho de Graesser et. al (2004) também destaca a importância da coesão na compreensão do texto, argumentando que a falta de coesão pode levar a dificuldades na compreensão. O Coh-Metrix avalia a coesão em várias dimensões, incluindo referência, conectividade lexical e estruturação de frases, e fornece uma pontuação geral de coesão para o texto.

O Coh-Metrix possui versões e variações apresentadas no decorrer dos anos e com diferentes pesquisas sobre a escrita, psicolinguística, linguística cognitiva e computacional. O Coh-Metrix é uma ferramenta focada na língua inglesa, sendo que suas métricas e avaliações são para esse idioma, assim em 2008 pesquisadores do Núcleo Interinstitucional de Linguística Computacional (NILC/USP) desenvolveram uma versão para o português brasileiro, chamado Coh-Metrix-Port. Hoje a terceira versão conta com 46 métricas de avaliação textual (Leal et. al., 2021).

Durante os anos de 2008 a 2021, diversas pesquisas foram desenvolvidas pelo grupo NILC, variações do software Coh-Metrix-Port foram desenvolvidas e a cada versão, novas métricas eram integradas e aprimoradas. Por fim, tem-se o software NILC-Metrix que fornece 200 métricas sobre o texto avaliado, divididas em 14 categorias, sendo elas: Medidas Descritivas, Simplicidade Textual, Coesão Referencial, Coesão Semântica, Medidas Psicolinguísticas, Diversidade Lexical, Conectivos, Léxico Temporal, Complexidade Sintática, Densidade de Padrões Sintáticos, Informações Morfossintáticas de Palavras, Informações Semânticas de Palavras, Frequência de Palavras e Índices de Leiturabilidade (Leal et. al., 2021).

Entre as métricas do NILC-Metrix, pode-se diferenciar algumas que se concentram nas características da sintaxe, enquanto outras se concentram na diversidade lexical, conectivos e fatores temporais. Leal et. al. (2021) informa que o sistema compreende aproximadamente 200 métricas as quais podem ser usadas para medir a legibilidade de um texto, incluindo o índice de legibilidade Brunet e a fórmula adaptada de Dale Chall.

Essa ferramenta tem sido empregada em várias investigações na área, exemplificada pelo projeto Adole-sendo (Leal et al., 2023). Esse projeto visava examinar os fatores biopsicossociais influenciando o desenvolvimento comportamental de adolescentes de 9 a 15 anos, com um foco específico em aspectos de maturação biológica. Dentro desse contexto, a idade cronológica dos participantes e as informações relacionadas foram utilizadas para desenvolver um modelo preditivo da complexidade nas narrativas dos adolescentes, fornecendo uma base sólida para o estudo.

Um outro uso do NILC-Metrix pode ser observado no corpus Leg2Kids (Leal et al., 2023), que desempenhou um papel fundamental na evolução dessa ferramenta. Esse corpus consistia em 36.413 legendas de filmes e séries dos gêneros Família e Animação em Português do Brasil, obtidas através da Open Subtitles em 2019. O processo de preparação do corpus envolveu a remoção de marcadores temporais de cada legenda, responsáveis por indicar o tempo de exibição na tela. Após essa etapa, o corpus foi dividido em frases e tokenizado com o uso da ferramenta NLTK, resultando em um total de 153.791.083 tokens e 452.312 tipos distintos.

Esses exemplos de pesquisa focam principalmente na identificação de marcadores de legibilidade em textos, com aplicações práticas na categorização de textos adequados para diferentes níveis educacionais ou perfis de leitores. Contudo, a utilização dessas métricas para investigar sua eficácia como indicadores de qualidade textual, especialmente no suporte à prática pedagógica, ainda não é uma abordagem comum. Esse aspecto constitui o foco principal deste trabalho.

2.4.2 Sobek

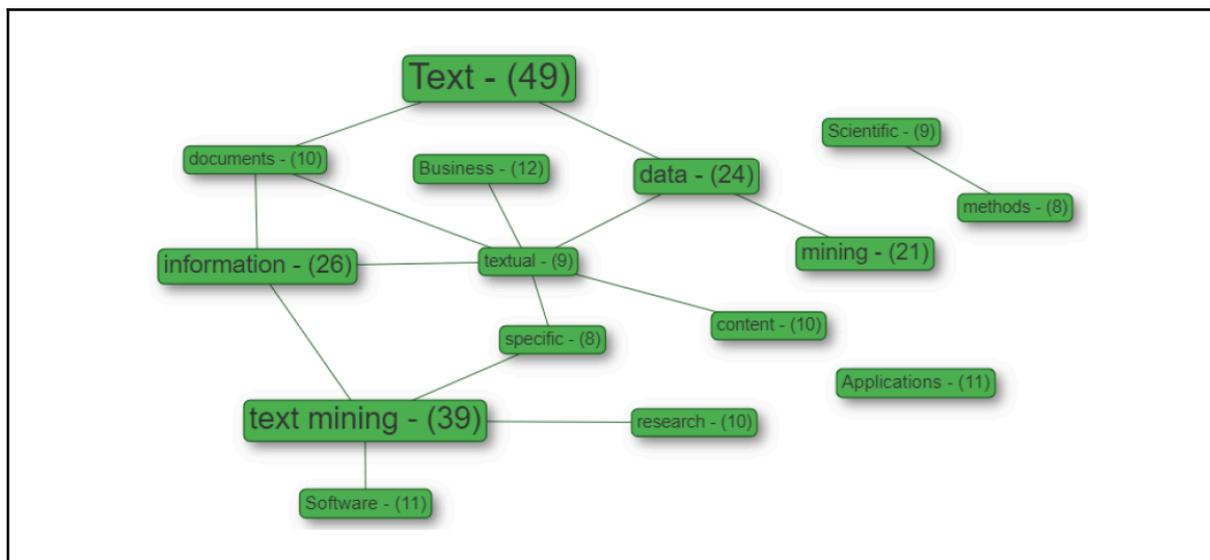
Softwares mineradores de texto desempenham papel importante em diversas áreas, como processamento de linguagem natural, extração e recuperação de informações, aprendizado supervisionado e não supervisionado, mineração de mídia social, entre outros. A aplicação e utilização de softwares de mineração textual para a análise, auxílio e avaliação da escrita não é novidade e é utilizado em trabalhos como Reategui et. al. (2011), Epstein e Reategui (2015) e Linck et. al. (2022).

O Sobek é um minerador de textos baseado em um algoritmo não supervisionado, com foco em aplicações educacionais. É mantido pelo GTech.Edu, grupo de pesquisa ligado à Universidade Federal do Rio Grande do Sul (UFRGS). Sua aplicação se dá através da mineração de textos individuais, sem a avaliação de bases de dados ou contextos desse texto. (Reategui et. al., 2022)

O Sobek inicia a mineração dividindo o texto em um conjunto de palavras mapeadas em relação aos seus termos e classificação gramatical e frequência ou repetição no texto. O minerador por padrão limita o número de termos relevantes em 20, porém permite ao usuário que esse informe o número de termos desejados ou selecione gráficos maiores, aumentando assim a quantidade de termos selecionados.

Reategui et. al. (2022) descreve o segundo passo do funcionamento do Sobek como sendo a identificação das relações entre os termos. Os pontos avaliados nessa etapa são a distância entre os termos, ou seja, quantas palavras existem entre eles e qual a taxa de frequência dessa relação.

Figura 1 - Gráfico de resultado do Sobek



Fonte: Reategui et. al. (2022)

O último passo do Sobek é a apresentação do resultado da mineração em forma gráfica. A figura 1 acima demonstra um exemplo de resultado de mineração de um texto retirado da Wikipédia sobre mineração de textos (Reategui et. al. 2022).

2.4.3 Ferramentas de feedback automatizado da escrita

Muitos trabalhos falam da avaliação da escrita acadêmica de um ponto de vista linguístico. O trabalho de Russell (2014), por exemplo, compara recursos linguísticos na escrita acadêmica de alunos avançados em inglês com estudantes universitários que utilizam inglês como primeira língua. Já o trabalho de Pulitzer-Ahles, Girolamo e Ghali (2020) investiga como duas pesquisas consideradas “boas” podem receber avaliações diferentes quando julgadas com base na sua redação em inglês, evidenciando um viés linguístico na revisão acadêmica.

Considerar trabalhos com foco e objetivos tão diferentes é relevante para esta tese do ponto de vista teórico, porque permite compreender categorias de análise que podem interessar à organização de um processo de avaliação por meio de recursos tecnológicos.

Do ponto de vista tecnológico, há muitas ferramentas que fazem a avaliação textual, como as ferramentas Icourse e Icourse+Pigai. Normalmente, tratam de análises mais pontuais, do ponto de vista gramatical/sintático. A ferramenta Icourse consiste em uma plataforma de gerenciamento de aprendizagem, utilizada para fornecer conteúdos e materiais de estudo aos alunos, já a ferramenta Pigai foi utilizada para fornecer aos alunos práticas de escrita e de correção.

No entanto, não são muitas as ferramentas que tratam de aspectos globais do texto, que buscam analisar aspectos textuais como coerência, precisão e complexidade. Essas, muitas vezes, estão voltadas à avaliação da escrita em língua adicional, como por exemplo o trabalho de Urmila (2021), que analisa a coerência e coesão na escrita de estudantes de inglês como segunda língua.

Nesse âmbito, cabe destacar que existem outras ferramentas que também fazem isso. Contudo, essas ferramentas não tratam de aspectos específicos de análise de projetos de pesquisa. A ferramenta desenvolvida por Bastiani (2022) tem esse foco, no entanto busca estabelecer uma comunicação direta com os estudantes, ao passo que a proposta desta tese é de construir uma ferramenta para apoiar o trabalho de avaliação de projetos pelo professor.

Sistemas computacionais podem ser treinados e/ou desenvolvidos com o objetivo de fornecer feedback automatizado e focado em alguns dos problemas comentados anteriormente. Por exemplo, as ferramentas Grammarly e Hemingway utilizam algoritmos para detecção de erros de gramática, ortografia, pontuação, estilo, entre outros.

A literatura acadêmica tem demonstrado como o uso de tecnologias na avaliação automatizada da escrita para fins formativos pode auxiliar os estudantes no aprimoramento de suas habilidades de produção textual (ZHAI e MA, 2022). Um estudo realizado por Tai et al. (2015), por exemplo, destaca a relevância do feedback para melhorar a qualidade da escrita de alunos universitários e enfatiza a importância de fornecer feedback personalizado, que leve em consideração as características individuais dos alunos. Segundo os autores, o uso de tecnologias pode auxiliar os professores na tarefa de fornecer feedback personalizado, garantindo a individualização do processo de aprendizagem e, conseqüentemente, melhorando a qualidade da escrita acadêmica. Além disso, os autores destacam a

necessidade de se buscar ferramentas que promovam a colaboração e a interação entre alunos e professores, permitindo uma troca de ideias e feedbacks construtivos.

Outro estudo realizado por Ching e Hsu (2013) destaca a importância do uso de tecnologias para o ensino de escrita acadêmica em ambientes virtuais de aprendizagem. Segundo os autores, o uso de tecnologias pode proporcionar um ambiente de aprendizagem mais efetivo e interativo, facilitando a colaboração e a construção coletiva do conhecimento.

De acordo com Bruton e Childers (2016), a utilização de ferramentas computacionais no processo de revisão e avaliação de trabalhos acadêmicos também pode auxiliar na identificação de plágio e na promoção da originalidade dos trabalhos.

Outra ferramenta que pode ser mencionada é o Feedback Studio, a qual permite aos professores fornecer feedbacks detalhados sobre a escrita dos alunos, destacando pontos fortes e fracos e oferecendo sugestões de melhoria. O Feedback Studio também permite que os alunos visualizem os feedbacks em tempo real, promovendo uma interação mais eficiente e eficaz entre alunos e professores.

O AcaWriter é uma ferramenta de análise de aprendizado que fornece feedback formativo sobre a redação acadêmica. O software é projetado para ajudar estudantes universitários a melhorar sua redação, fornecendo feedback automatizado sobre seus textos.

O AcaWriter usa algoritmos de processamento de linguagem natural para analisar o texto e identificar determinados padrões, como declarar uma questão de pesquisa, resumir pesquisas anteriores ou fornecer evidências para apoiar uma afirmação. Segundo Knight et. al. (2020), o AcaWriter compara o texto do aluno com um banco de dados de expressões comumente usados na escrita acadêmica. A partir dessa análise, o software fornece feedback sobre a qualidade da redação do aluno, destacando áreas a serem melhoradas.

Outra ferramenta importante a ser mencionada é o plug-in denominado MTA (BASTIANI, 2022). Esse plug-in, desenvolvido para o editor de textos on-line Google Docs, visa auxiliar o estudante durante o processo de escrita de projetos de pesquisa. Essa ferramenta, através de dicas e feedbacks positivos e corretivos

fornece ao estudante orientações para auxiliá-lo na identificação de diferentes problemas relativos à escrita de projetos de pesquisa, tais como formulação da questão de pesquisa, coerência entre título/objetivo/referencial teórico e descrição de aspectos metodológicos do trabalho.

Além de ferramentas que permitam ou auxiliem o processo de comunicação e transmissão do feedback entre professores e alunos, é possível utilizar ferramentas que avaliem pontos específicos dos trabalhos e produzam feedbacks automatizados, focados em auxiliar o professor na identificação de pontos e características pré-definidas dos textos dos alunos.

Esse modelo de ferramenta computacional pode auxiliar o processo de avaliação do professor, reduzindo sua carga de trabalho, diminuindo o tempo de correção e feedback ao aluno, além de tornar a avaliação mais precisa e justa para os alunos (Cavalcanti, 2021). O objetivo de uso de ferramentas de feedback automatizado não é remover o professor da tarefa de avaliação dos trabalhos, mas sim auxiliar esse avaliador, fornecendo uma visão inicial do projeto e permitindo que o professor identifique rapidamente os principais problemas e áreas que podem ser melhoradas, podendo assim, concentrar-se em fornecer um feedback mais detalhado e personalizado aos alunos em áreas específicas de seus trabalhos.

Cavalcanti (2021), ainda comenta a dificuldade encontrada em fornecer feedbacks aos alunos em turmas muito numerosas, podendo assim acontecer a implementação de sistemas de feedback automatizado, com a finalidade de auxiliar o professor nessa tarefa.

Assim, é importante que as ferramentas de feedback automatizado sejam desenvolvidas com base em critérios claros e objetivos de avaliação. De acordo com Wang e Han (2022), as ferramentas de feedback automatizado podem ser eficazes apenas se forem capazes de avaliar adequadamente os critérios de avaliação definidos pelos professores. Os autores destacaram a importância de envolver os professores na definição desses critérios e na validação das ferramentas de feedback automatizado antes de sua implementação.

No trabalho de Liu et. al. (2017), os professores foram envolvidos na configuração da ferramenta de feedback. Nesse trabalho foram analisadas redações

de estudantes do segundo ano do College of International Studies at Southwest University, que, após serem avaliadas e classificadas por professores especialistas, foram integradas ao software Coh-Metrix onde foram mensuradas entre três métodos. Por fim, novos alunos foram submetidos a comparativos, onde alguns receberam feedback personalizado dos professores e outros receberam feedbacks automáticos gerados pelo sistema utilizado. Como conclusão, os autores identificaram que o feedback indireto, ou seja, o automatizado, produziu mais correções e revisões ao texto.

Já o trabalho de Cavalcanti (2021) apresenta uma revisão sistemática sobre a geração automática de feedback em sistemas de apoio à aprendizagem. Os autores obtiveram resultados mostrando que 82,53% dos estudos analisados não evidenciaram maior eficácia no feedback manual em detrimento do feedback automatizado. Contudo, como já argumentado anteriormente, há aspectos problemáticos no uso exclusivo de feedback automatizado, como a dificuldade em avaliar habilidades de mais alta ordem, como argumentação e pensamento crítico, bem como prover feedback personalizado considerando nuances algumas vezes sutis relacionadas a cada estudante (Wang e Han, 2022, Gallien e Early 2005).

Assim como Wang e Han (2022) argumentam que as ferramentas de feedback automático devem ser adequadas ao formato e modelo de texto a ser analisado, é importante atentar também a forma de apresentação do feedback ao destinatário. Na maioria dos trabalhos analisados, o destinatário foco da ferramenta é o aluno, assim, dashboards e ferramentas visuais nem sempre são compreendidas, o que atrapalha a compreensão do feedback recebido. Cavalcanti (2021) expõe em sua revisão que alguns dos artigos analisados relatam a insatisfação dos alunos com os feedbacks recebidos, onde foram listados argumentos como a brevidade, a polaridade sempre negativa e a complexidade em interpretar o feedback recebido.

Devido a isso, é importante observar as qualidades e limitações de cada feedback para que uma ferramenta de feedback automatizado seja capaz de contribuir de forma satisfatória aos objetivos propostos, acelerar o processo de correção de trabalhos ou melhorar a escrita acadêmica dos alunos.

Nesse sentido, mostra-se vantajoso associar o feedback automatizado àquele do professor, o qual conta com conhecimentos específicos acerca dos objetivos, estágio e limitações da disciplina, turma e alunos. O feedback do professor também tende a ser mais abrangente que o feedback automatizado, uma vez que pode fazer inferências e sugestões sobre pontos ou temas que a ferramenta computacional não tenha sido programada para abordar. (Thi e Nikolov, 2022 e Woodworth, 2022).

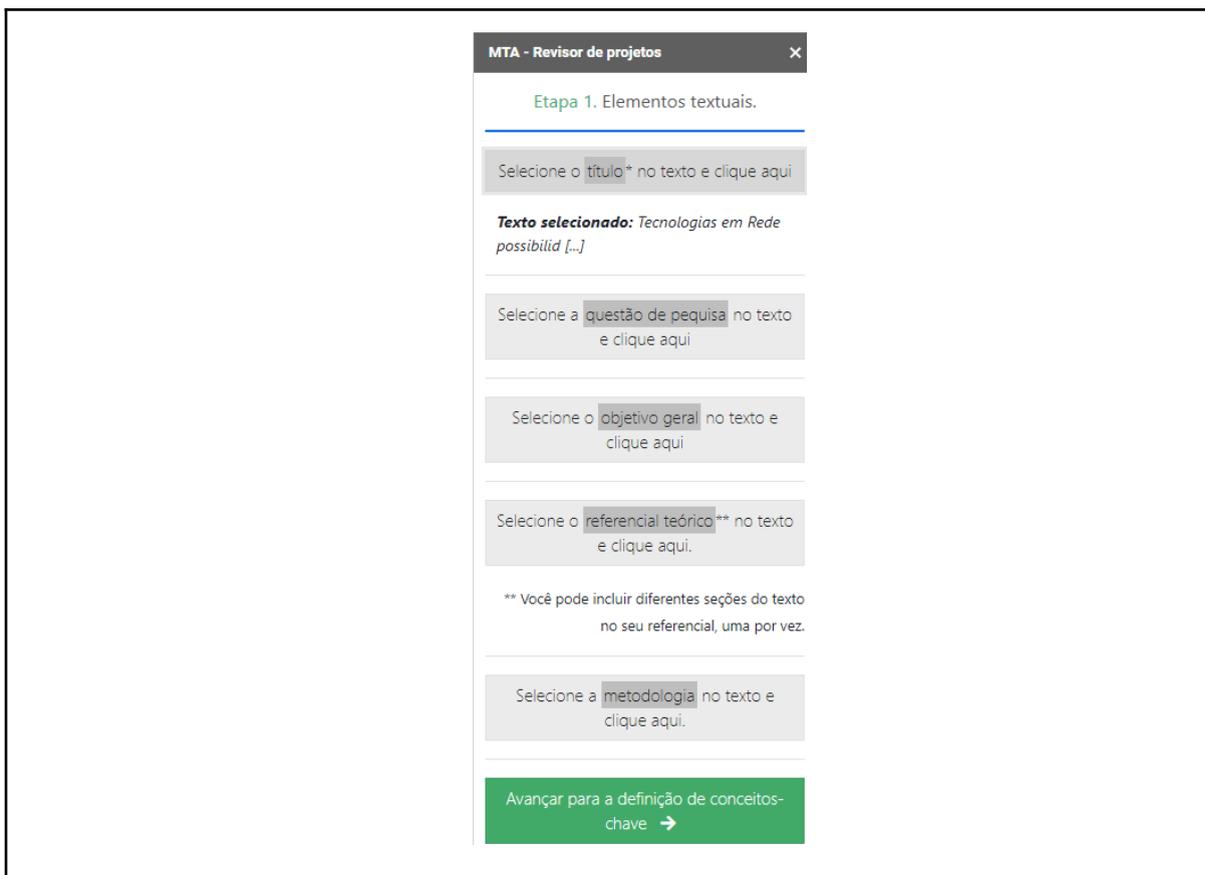
Dessa forma, reforça-se a escolha da abordagem proposta nesta tese em que a devolução do feedback ao aluno é feita pelo professor, buscando reduzir o tempo de avaliação e auxiliando-o na tarefa de análise dos trabalhos, permitindo que o professor ajuste e adeque o feedback conforme as necessidades e realidades de cada aluno.

2.4.4 O plug-in MTA

Nesta pesquisa, a proposta de tecnologia desenvolvida foi baseada na continuação/evolução do trabalho de Bastiani (2022). Por esse motivo, o plug-in MTA é aqui apresentado com um maior nível de detalhe. Originalmente, Bastiani (2022) desenvolveu um plug-in para o editor de textos on-line Google Docs, chamado MTA. Esse plug-in tem como objetivo auxiliar o aluno na revisão e melhoria escrita de seus projetos de pesquisa através de dicas e feedbacks formativos.

Para a utilização da ferramenta de Bastiani (2022), o aluno deve selecionar pontos chave do texto de seu projeto e atribuí-los ao plug-in. A figura 2 abaixo demonstra a tela de seleção de elementos textuais da referida ferramenta.

Figura 2 - Campos do plug-in MTA

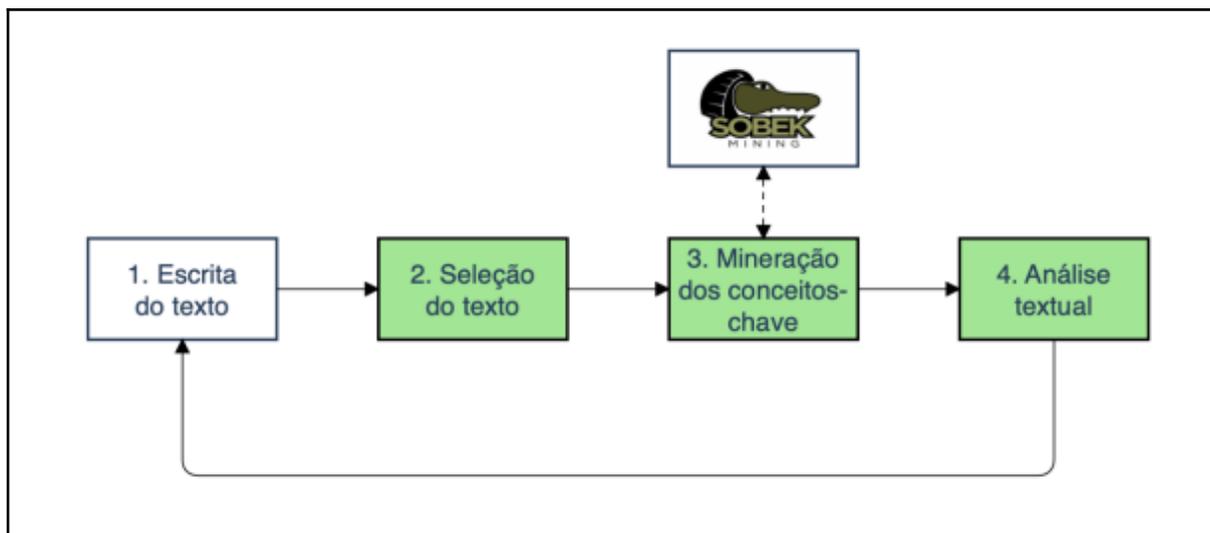


Fonte: Plug-in MTA de Bastiani (2022)

Pode-se notar através da figura 2, que a ferramenta avalia e revisa pontos como o título, a questão de pesquisa, o objetivo geral, o referencial teórico e a metodologia do trabalho. Segundo o autor, o relacionamento entre esses pontos é crucial para a identificação de falhas estruturantes em projetos de pesquisa, sendo o plug-in MTA capaz de avaliar características específicas de cada elemento, como por exemplo a existência de uma pergunta na questão de pesquisa.

Durante a utilização da ferramenta, são extraídos os conceitos-chave do texto em análise. Essa extração é feita via mineração de texto, com a utilização do software Sobek, mencionado na seção acima.

Figura 3 - Fluxograma do plug-in MTA



Fonte: Bastiani (2022) p.64

A figura 3 acima descreve o fluxo de funcionamento da ferramenta. Após a escrita do texto, parte onde o plug-in não tem ação, o usuário deverá selecionar o texto correspondente a cada ponto avaliativo do plug-in. Ao avançar no uso, serão minerados os conceitos chave através do software Sobek, finalizando com a análise textual que retorna ao usuário recomendações de melhoria e feedbacks.

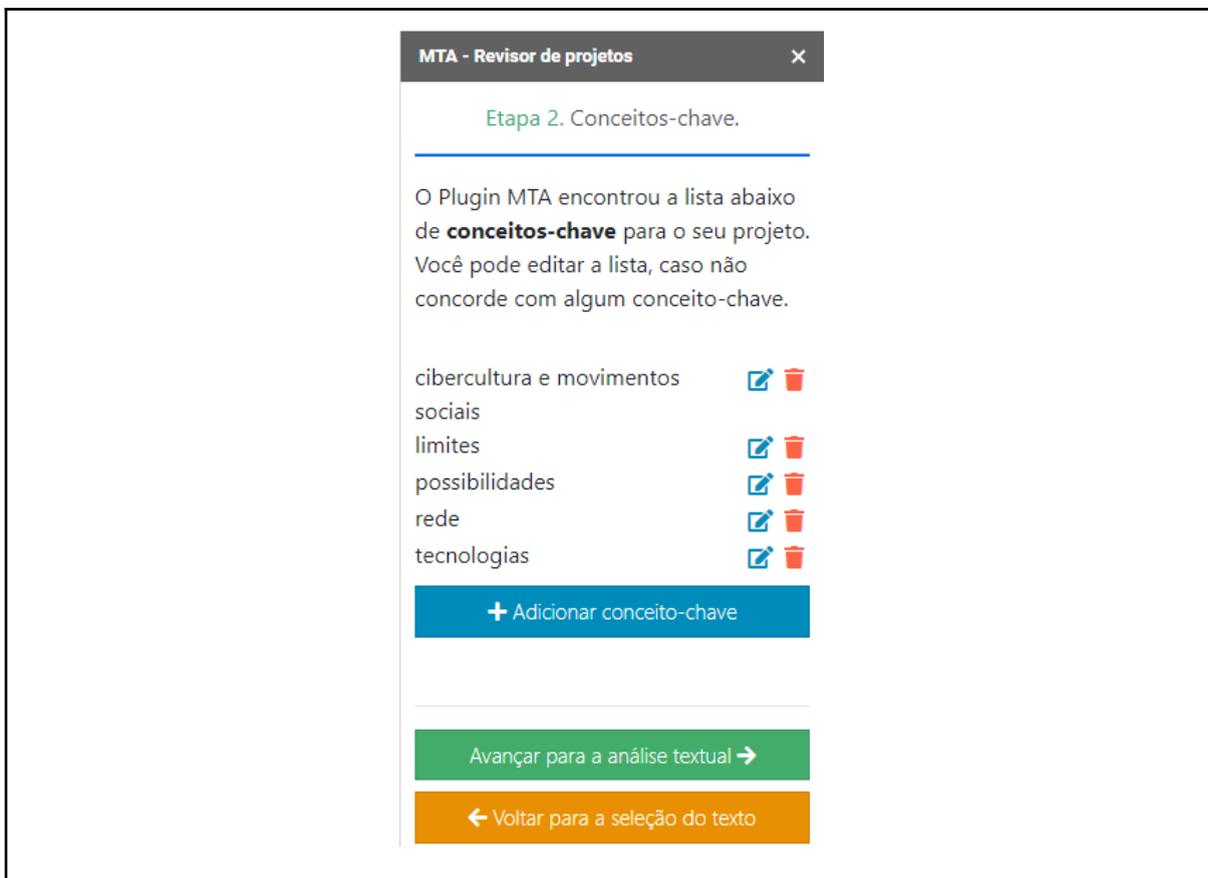
Como terceiro passo do fluxograma de Bastiani (2022), os conceitos chave extraídos pelo minerador de textos Sobek serão apresentados ao usuário. Esses elementos estruturantes do projeto de pesquisa são extraídos conforme o título, a questão de pesquisa e/ou objetivo geral selecionados na etapa anterior de uso do MTA.

Segundo Reategui (2020), é importante a definição desses conceitos, uma vez que existe a tendência do estudante em utilizar questões de pesquisa amplas demais. Ainda segundo Reategui (2020), a boa condução e escolha do título do trabalho é fundamental, uma vez que ele é o primeiro contato do público com o conteúdo da pesquisa. Dessa forma, o título deve expressar de maneira clara o conteúdo, assim como o objetivo deve esclarecer o propósito e a questão de pesquisa que será respondida.

A figura 4 abaixo demonstra a maneira como os conceitos-chave são apresentados pelo plug-in MTA, ressaltando-se a permissão de inserção, alteração e exclusão dos conceitos extraídos. Bastiani (2022) explica que essas permissões são

necessárias pois entende-se que uma ferramenta automática de mineração de textos pode apresentar elementos imprecisos ou com os quais o autor do texto não concorde totalmente.

Figura 4 - Extração de conceitos chave plug-in MTA



Fonte: Plug-in MTA de Bastiani (2022)

O quarto passo do plug-in MTA consiste na análise dos elementos textuais e o fornecimento de feedback ao usuário. As figuras 5 e 6 abaixo exemplificam esse funcionamento.

Figura 5 - Exemplo de texto avaliado pelo MTA

Título	Objetivo Geral	Questão de pesquisa
<i>Desenvolvimento e validação de uma ferramenta computacional baseada na estratégia de outlining para apoio à escrita de projetos de pesquisa.</i>	<i>como uma ferramenta computacional, apoiada pela estratégia de outlining pode qualificar o processo de escrita.</i>	<i>uma ferramenta computacional apoiada pela estratégia de outlining</i>

Fonte: Bastiani (2022) p.69

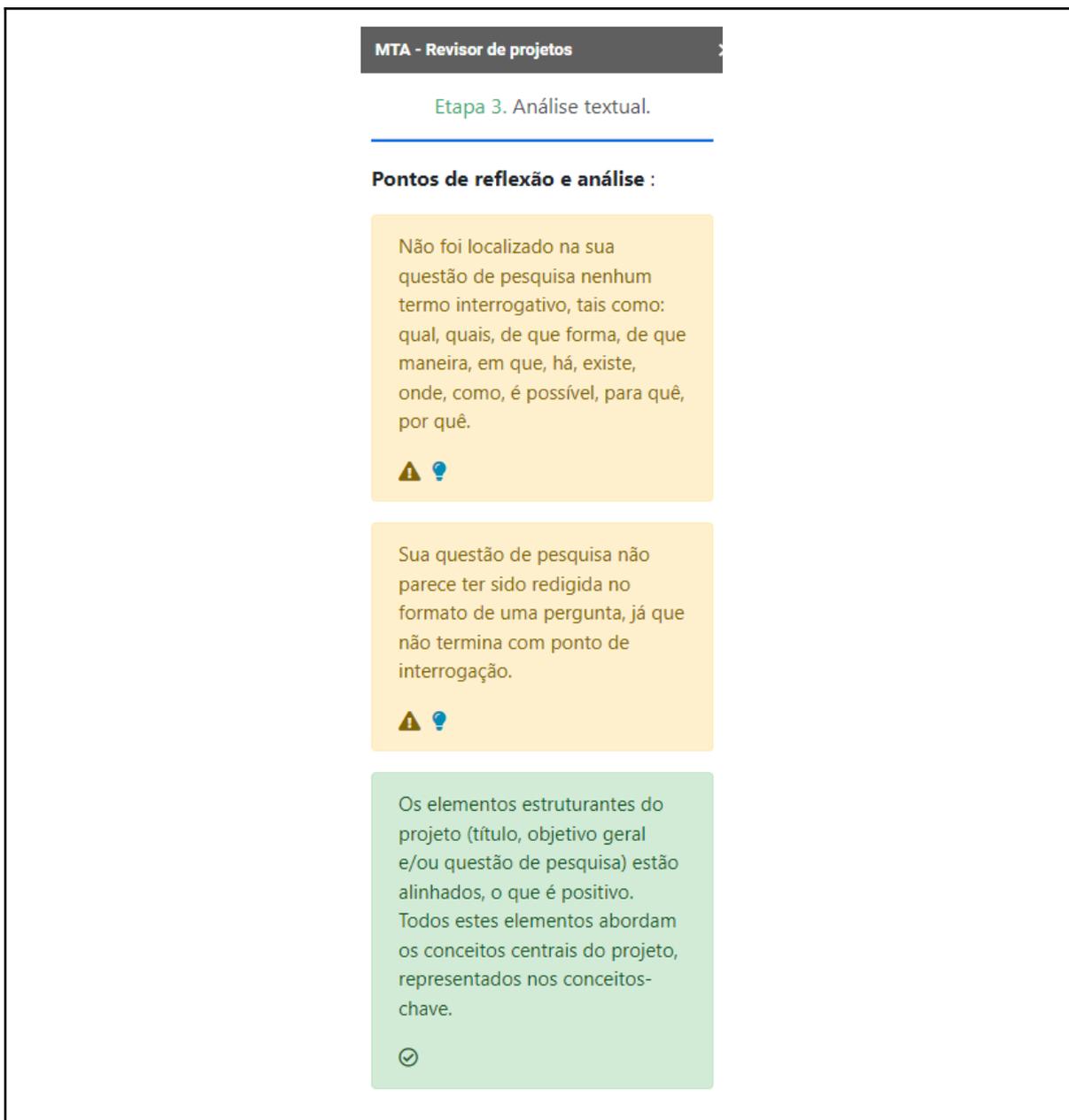
Na figura 5 acima se veem exemplos de textos extraídos e utilizados no plug-in MTA. Cada um dos elementos, título, objetivo geral e questão de pesquisa foram separados e identificados nos campos destinados a cada um dentro do plug-in. Com isso, a ferramenta, ao analisar cada elemento, identificará que, por exemplo, o objetivo geral não está condizente com algumas orientações acadêmicas. Segundo Praça (2015), o objetivo geral deve indicar uma ação do pesquisador, normalmente representada por um verbo no infinitivo, como por exemplo, aprender, analisar, entender, estudar, entre outros.

Gil (2015) por sua vez argumenta que a questão de pesquisa deve apresentar um questionamento, o que obriga que essa seja formulada como uma pergunta.

A figura 6 abaixo apresenta um exemplo de feedback da ferramenta MTA. Esse feedback é dividido e fornecido em duas categorias, sendo uma de feedback de reflexão e outra de feedback de reforço positivo. Bastiani (2022) informa que vinculou o retorno da ferramenta desta forma pois autores como Hayes et. al. (1987) argumentam sobre a dificuldade que os escritores possuem em detectar suas falhas durante a revisão do próprio texto.

Assim, o autor reforça a divisão em duas categorias de feedback, tendo o reflexivo o objetivo de estimular o aluno a melhorar seu texto. Já o feedback positivo visa a orientar o aluno, reforçando sua confiança durante a escrita de projetos.

Figura 6 - Feedback do plug-in MTA



Fonte: Plug-in MTA de Bastiani (2022)

Um ponto negativo desta ferramenta consiste na sua limitação a documentos da plataforma Google Docs. Uma vez que foi desenvolvida através da Google Apps Script, essa ferramenta não é capaz de avaliar e fornecer feedbacks para projetos de pesquisa escritos em outros editores, como por exemplo o Microsoft Word.

Outro ponto a ser considerado consiste no público alvo do plug-in MTA. Esse plug-in foi desenvolvido para auxiliar o aluno durante o processo de escrita de

projetos de pesquisa, ou seja, para fazer a avaliação de projeto, é necessário que o usuário, seja professor ou aluno, abra o documento, selecione os pontos-chave e execute o plug-in dentro do projeto. Essa característica torna o MTA trabalhoso para ser utilizado como ferramenta auxiliar ao professor, não fornecendo mecanismos para avaliação de grande escala nem contendo meios de verificação entre versões do mesmo projeto.

3 METODOLOGIA DA PESQUISA E RESULTADOS

Nesta seção será apresentada a metodologia adotada para desenvolvimento da pesquisa, dividida em 3 etapas:

- **Etapa 1:** Desenvolvimento da ferramenta para avaliação de projetos de pesquisa;
- **Etapa 2:** Análise de métricas de complexidade e estrutura textual para dois conjuntos de textos distintos, a saber: redações de vestibular e projetos de pesquisa em nível de graduação. O objetivo desses estudos foi verificar eventuais relações entre essas métricas e avaliações realizadas por professores e revisores, fornecendo insumos para a implementação de funções de análise da ferramenta apresentada na seção anterior;
- **Etapa 3:** Estudo de casos múltiplos, com utilização da ferramenta por professores de disciplinas de metodologia da pesquisa.

Os estudos realizados na Etapa 2 tiveram abordagem fundamentalmente quantitativa. Já a Etapa 3 baseou-se na modalidade de estudo de casos múltiplos, do qual participaram 3 professores de uma mesma disciplina relacionada à metodologia da Pesquisa em Educação. O foco dos estudos foi avaliar em que medida a ferramenta desenvolvida pôde auxiliar os professores na análise e avaliação dos trabalhos dos estudantes a partir da geração de feedbacks automáticos. Seguindo as definições de Gil (2008), os estudos de casos múltiplos foram desenvolvidos de maneira detalhada, buscando-se acompanhar de maneira mais próxima as ações realizadas pelos professores no uso da ferramenta. A análise dos dados da Etapa 3 se deu seguindo abordagem qualitativa e quantitativa. Os dados qualitativos ficaram a cargo das entrevistas com os professores participantes da pesquisa e os dados quantitativos foram obtidos através de coleta realizada por meio da própria ferramenta.

Cada uma das 3 etapas da pesquisa, acompanhada dos respectivos resultados, é descrita a seguir.

3.1 Etapa 1: Desenvolvimento da ferramenta

3.1.1. Aspectos relacionados à implementação

Para responder ao objetivo geral da pesquisa, foi desenvolvida uma ferramenta computacional para avaliação de projetos de pesquisa voltada para o professor, a qual recebeu o nome de ATHOS. Ela conta com as funções de avaliação e geração de feedbacks automatizados que visam a auxiliar o professor na correção e identificação de elementos centrais dos projetos de seus alunos, como identificação do título, objetivo e palavras-chave. Tal identificação teve como objetivo agilizar o processo de avaliação dos trabalhos e a devolução de feedbacks. A ferramenta faz uso de sistemas externos, como ferramentas de processamento textual e PLN como o NILC-Metrix, apresentado na seção 2.4.1, ferramentas de IA, sendo o CHAT GPT 3.5 a escolhida para a função e também métricas do Plugin-MTA, apresentado em detalhe na seção 2.4.4, desenvolvido na pesquisa de doutorado de Bastiani (2022). A figura 7 ilustra o fluxograma de funcionamento da ferramenta e apresenta suas funcionalidades. As caixas verdes representam ações automatizadas da ferramenta, enquanto as caixas azuis representam as ações realizadas pelos professores.

A ferramenta Athos conta com um módulo de conexão com a plataforma Moodle para coleta dos trabalhos dos alunos. Esse módulo faz uso de um token de identificação do professor para a conexão com o servidor, permitindo a identificação da turma e atividade que se deseja a coleta dos trabalhos.

A preparação dos arquivos envolve a conversão do formato pdf em docx, quando necessário, além da remoção de tabelas e figuras, mantendo apenas a parte textual do trabalho. Ainda são removidas páginas comuns com capa, apresentação e sumário, além de seções de referências bibliográficas, apêndices e anexos. Ainda nesse ponto da ferramenta Athos, as páginas de capa, resumo e sumário são enviadas a ferramentas de IA para identificação e extração do título do trabalho.

A conversão de formatos é necessária uma vez que a extração textual de um arquivo pdf, não mantém as estruturas textuais de parágrafos e quebras de linha, o que resulta em valores incorretos no processamento da ferramenta NILC-Metrix. A

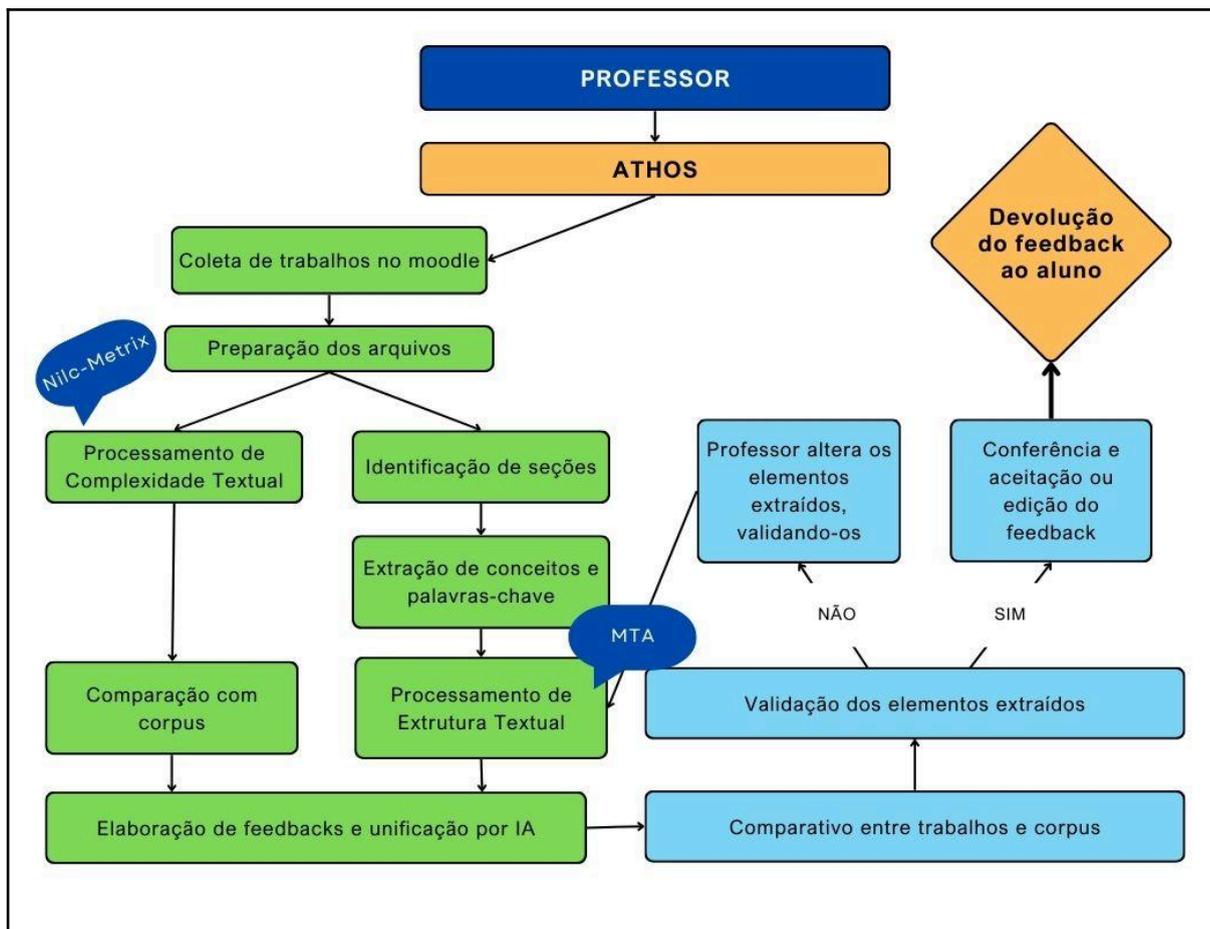
finalização da limpeza foi efetuada com a remoção das tabelas e quadros de texto dos arquivos via algoritmo e técnicas de PLN em linguagem Python.

O módulo de Processamento da Complexidade Textual consiste no cálculo das métricas da ferramenta NILC-Metrix e de outras métricas relacionadas, como tamanho das seções. Para execução desse módulo, fez-se necessária a instalação e implementação do referido sistema. Esse procedimento foi efetuado conforme os passos disponibilizados pelo grupo NILC na plataforma GitHub, disponível em: <https://github.com/nilc-nlp/nilcmatrix>.

Devido a limitações computacionais e/ou das tecnologias utilizadas pela ferramenta, fez-se necessário dividir os arquivos em blocos para realização do processamento. Observa-se que o próprio grupo NILC disponibiliza a ferramenta com a limitação de 2.000 palavras por processamento. A ferramenta está disponível em: <http://fw.nilc.icmc.usp.br:23380/nilcmatrix>.

Os resultados das métricas geradas pela ferramenta NILC-Metrix foram padronizados em quantidade de desvio-padrão da média gerada pelos estudos desenvolvidos previamente desenvolvidos. Os estudos estão descritos na sequência e também com mais detalhes na seção 4.1.

Figura 7 - Diagrama de funcionamento da ferramenta Athos



Fonte: do autor.

A identificação das seções é feita através de algoritmos e técnicas de PLN com objetivo de identificar as seções do trabalho, como introdução, referencial e metodologia, separando e preparando-as para os próximos passos.

Cada seção identificada é submetida à extração de elementos e palavras-chave. Essa extração é efetuada através da ferramenta de IA generativa ChatGPT na versão 3.5. O objetivo dessa extração é permitir o cálculo da semelhança e a ocorrência dos termos chave em cada uma das seções, sendo esse um dos métodos de determinação do grau de coerência entre as seções.

O módulo de Processamento de Estrutura Textual é composto em base pelo plugin MTA, de Bastiani (2022), porém, esse plugin teve modificações em seu funcionamento com o objetivo de operar de forma automatizada, não sendo mais necessário que o usuário selecione as seções do trabalho manualmente. Para isso, fez-se necessária a reestruturação do código para a linguagem JAVA, permitindo

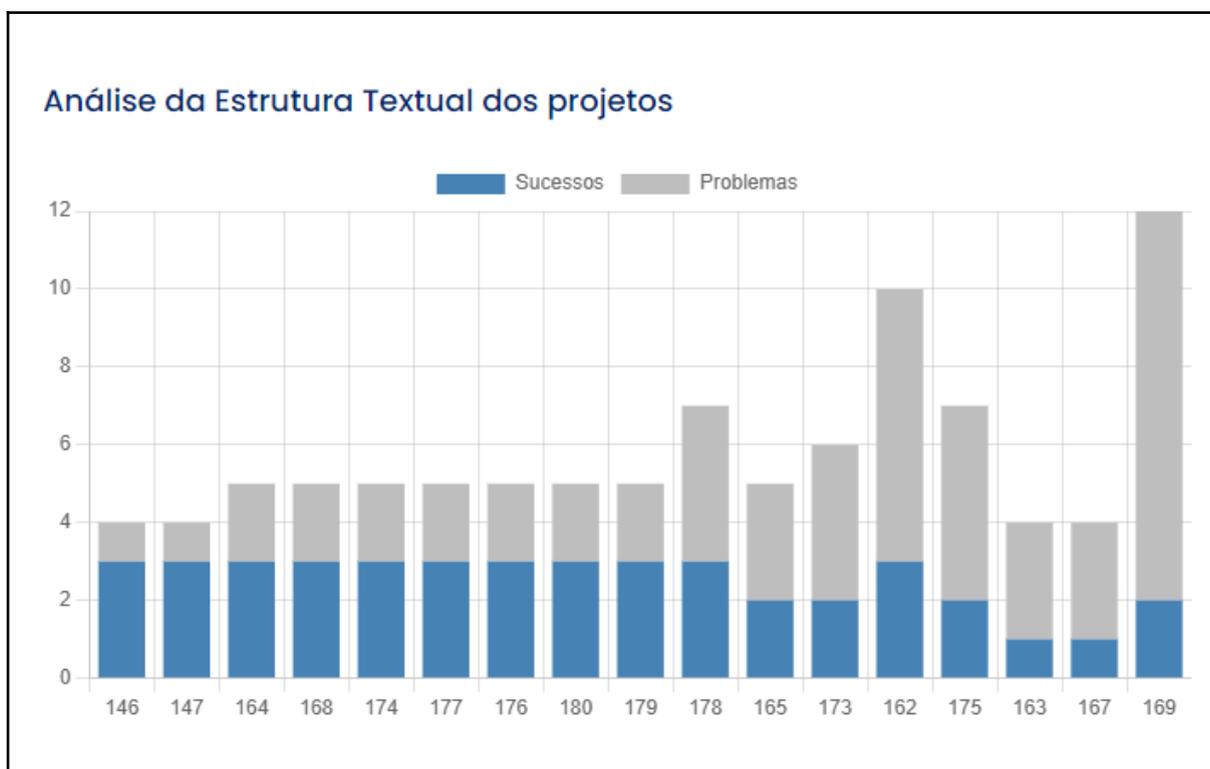
assim o processamento dos arquivos fora da plataforma Google Docs. Os pontos avaliativos do plugin, já validados na tese de Bastiani (2022) seguiram inalterados.

Tanto o módulo de Processamento de Complexidade Textual quanto o módulo de Processamento de Estrutura Textual acionam respostas geradoras de feedbacks acerca do trabalho analisado. Essas respostas são retornadas em forma de texto específico para cada uma delas. Esses fragmentos são então enviados à API do Chat GPT para a conversão dessa sequência de feedbacks em um texto coeso e unificado.

Após a conclusão do processamento automatizado, o trabalho do aluno torna-se disponível ao professor para comparação, validação e avaliação.

A ferramenta Athos fornece ao professor dois gráficos para comparação do trabalho do aluno com seus pares e base de dados construída. O primeiro gráfico refere-se às métricas de Estrutura Textual, figura 8, já as métricas de Complexidade Textual são apresentadas na figura 9.

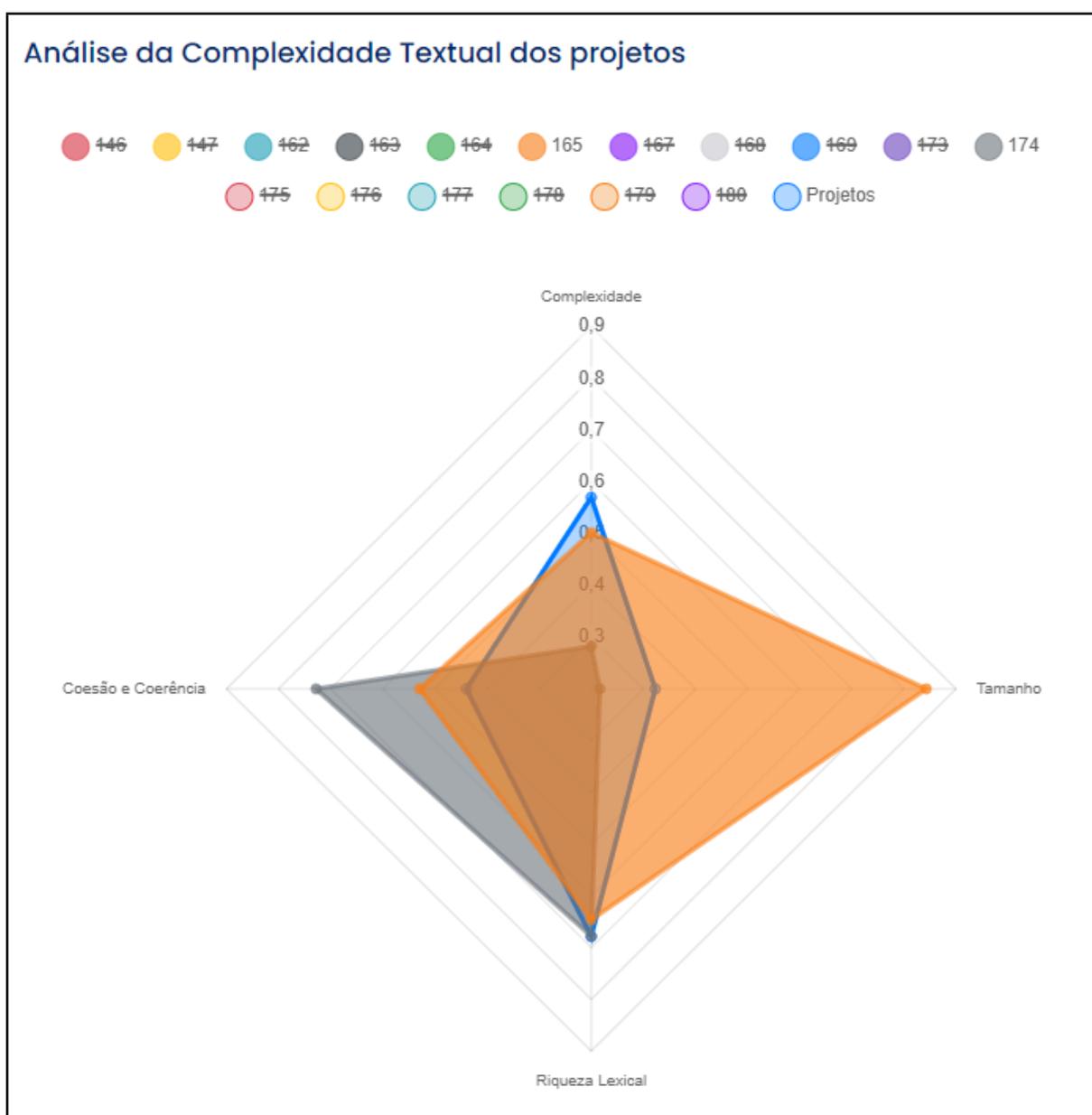
Figura 8 - Gráfico da análise estrutural dos projetos



Fonte: Sistema Athos

No gráfico da figura 8 acima, é possível vislumbrar os trabalhos enviados pelos alunos e sua avaliação quanto às métricas de estrutura textual, sendo apresentadas as quantias de retornos de “sucesso” e “problemas” representados pelas ocorrências das métricas e conseqüentemente refletidos em feedbacks para o professor. Com esse gráfico o professor pode previamente já identificar e concentrar-se nos trabalhos mais problemáticos.

Figura 9 - Gráfico de complexidade textual dos projetos



Fonte: Sistema Athos

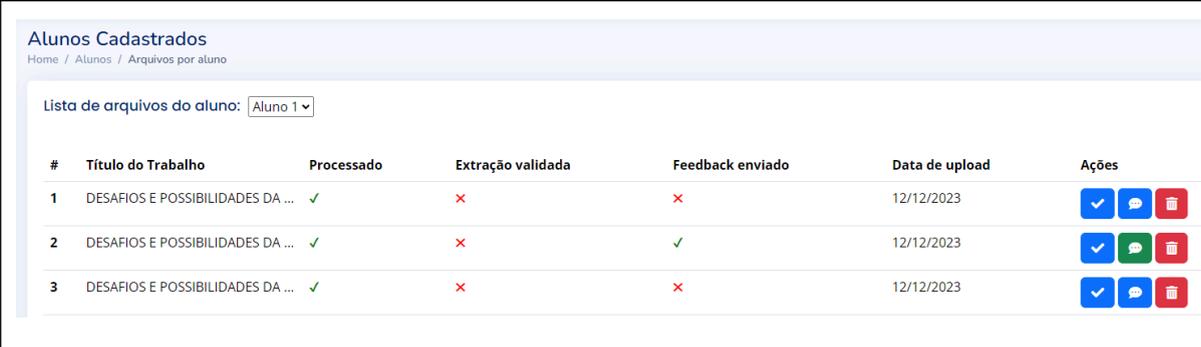
Na figura 9, no gráfico de complexidade textual, é fornecido ao professor participante uma visualização de como os trabalhos analisados se posicionam

quanto às métricas de complexidade textual, permitindo uma comparação com os trabalhos dos demais alunos e com a média construída pelos estudos anteriores.

Ambos os gráficos das figuras 8 e 9 foram desenvolvidos para permitir ao professor uma visualização facilitada de como os trabalhos se classificam em termos de estrutura e complexidade, permitindo que este professor possua uma compreensão prévia de onde o trabalho do aluno pode demandar maior atenção durante a correção.

Além dos gráficos apresentados pela ferramenta, o módulo do professor ainda permite que ele selecione com qual projeto ou aluno ele deseja trabalhar no momento, sendo possível listar todos os envios de determinado aluno, ficando assim, de fácil acesso aos últimos envios e trabalhos ainda não avaliados.

Figura 10 - Tela de atividades enviadas pelo aluno.



#	Título do Trabalho	Processado	Extração validada	Feedback enviado	Data de upload	Ações
1	DESAFIOS E POSSIBILIDADES DA ...	✓	✗	✗	12/12/2023	✓, 🗨️, 🗑️
2	DESAFIOS E POSSIBILIDADES DA ...	✓	✗	✓	12/12/2023	✓, 🗨️, 🗑️
3	DESAFIOS E POSSIBILIDADES DA ...	✓	✗	✗	12/12/2023	✓, 🗨️, 🗑️

Fonte: do autor.

Na figura 10, é apresentada a tela do sistema que mostra as atividades enviadas por um aluno. Essa parte do sistema permite ao professor ver a data do upload do arquivo, o título do trabalho, se a extração de dados já foi validada e se o feedback já foi enviado ao aluno. Também é possível acessar a página de validação de extrações, a página de feedback e deletar o envio. É importante destacar que é necessário que a atividade na plataforma Moodle seja do tipo recebimento de arquivo e não escrita na própria plataforma.

Figura 11 - Tela de atividades do professor.

Trabalhos Enviados
Home / Trabalhos / Trabalhos Enviados

Lista de trabalhos enviados

#	Aluno	Título do trabalho	Processado	Extração validada	Feedback enviado	Data de upload	Ações
1		DESAFIOS E POSSIBILIDADES DA EDUCAÇÃO BÁSICA C...	✓	✗	✗	13/12/2023	✓, 🗨️, 🗑️
2		INDUSTRIA CULTURAL E EDUCAÇÃO NOS TEMPOS DO ...	✓	✗	✗	13/12/2023	✓, 🗨️, 🗑️
3		COMO ATIVIDADES PRÁTICAS NO ENSINO DA QUÍMICA ...	✓	✗	✗	13/12/2023	✓, 🗨️, 🗑️
4		OS TICs NA EDUCAÇÃO	✓	✗	✗	13/12/2023	✓, 🗨️, 🗑️
5		Como o uso de inteligência artificial impacta a vida dos ...	✓	✗	✗	13/12/2023	✓, 🗨️, 🗑️
6		A importância da educação sexual como meio de preve...	✓	✗	✗	13/12/2023	✓, 🗨️, 🗑️
7		A interdisciplinaridade como abordagem para estimula...	✓	✗	✗	13/12/2023	✓, 🗨️, 🗑️

Fonte: Sistema Athos

A figura 11 apresenta a tela de relatório de atividades pendentes de correção. Nela é possível acessar a validação dos elementos extraídos e os feedbacks gerados pelo sistema Athos.

Ao acessar um trabalho específico e iniciar sua avaliação, a ferramenta Athos exige do professor que esse valide os elementos extraídos de maneira automatizada. Essa validação é necessária para mitigar possíveis problemas gerados pelas ferramentas de IA utilizadas no decorrer do processo.

Caso o professor julgue que os elementos extraídos não estão de acordo com o trabalho, ele tem a oportunidade de alterá-los, o que levará a ferramenta Athos a re-executar o módulo de Processamento de Estrutura Textual, podendo gerar assim novos feedbacks e por consequência, um novo texto. Esse processo pode ser repetido quantas vezes o professor desejar, desde que efetue alterações nos elementos extraídos.

Figura 12 - Tela de validação de extrações.

The screenshot displays the 'Tela de validação de extrações' interface. On the left, a document viewer shows a page with the heading '1 INTRODUÇÃO' and several paragraphs of text discussing the impact of Cultural Industry on education in Brazil. On the right, a validation panel contains four sections, each with a text input field and a rich text editor toolbar:

- Titulo do trabalho:** The input field contains 'INDUSTRIA CULTURAL E EDUCAÇÃO NOS TEMPOS DO TIKTOK'.
- Objetivo geral:** The input field contains 'Estudar o impacto da Indústria Cultural na educação de adolescentes no Brasil'. This section is highlighted with a red border.
- Questão de pesquisa:** The input field contains 'Quais são os efeitos da Indústria Cultural na educação? Como a literatura brasileira tem se posicionado e o que se tem produzido até o momento?'. This section is also highlighted with a red border.
- Palavras-chave:** The input field contains 'Indústria Cultural, Educação, Adolescente'.

Fonte: Sistema Athos

Na figura 12 é apresentada a tela de validação dos dados extraídos com o uso da plataforma Chat GPT. Nela, o professor pode acessar o trabalho do aluno e os dados extraídos, verificando sua adequação. As caixas sinalizadas por bordas vermelhas indicam que o texto extraído não foi encontrado, de maneira exatamente igual, no corpo do texto do aluno. Isso se dá por erros de escrita do aluno, ajustes na frase ou inferências efetuadas pela IA.

Quando o professor se dá por satisfeito com os elementos, título, objetivo e palavras-chave, pode visualizar o texto de feedback gerado pela ferramenta Athos. Esse texto pode ser editado conforme o professor julgar necessário, incluindo, excluindo ou aceitando o feedback fornecido, podendo devolver esse feedback ao aluno.

origem os trabalhos solicitados pelo professor no decorrer da disciplina e postados pelos alunos na plataforma Moodle.

Elencaram-se as variáveis de coleta como sendo:

- O grau de alterações no feedback efetuadas pelo professor.
- O número de feedbacks automáticos aceitos na íntegra.
- O número de feedbacks corretivos ou positivos gerados pela ferramenta. Aqui espera-se encontrar a complexidade textual do aluno/turma.
- O tempo despendido pelo professor para a utilização da ferramenta e o quanto isso impacta no tempo de trabalho.

3.1.2. Resultados sobre o desenvolvimento da ferramenta

Como mencionado anteriormente, a ferramenta desenvolvida neste trabalho utilizou como base o sistema desenvolvido em Bastiani (2022). Foram utilizadas as métricas de identificação de pontos-chave já validadas no referido trabalho e associadas com a ferramenta de IA generativa Chat GPT, além de técnicas de PLN para extração dos principais elementos e adequação e melhoria dessas para as particularidades do público alvo deste trabalho. A estes critérios, acresceram-se funções de análise de produção textual usando a implementação da ferramenta NILC-Metrix, também apresentada anteriormente.

Para a avaliação dos trabalhos de corpora, e conseqüentemente a geração dos feedbacks pela ferramenta Athos, foram utilizadas as métricas descritas na tabela 1. As métricas estão divididas em dois tipos de avaliações, sendo essas Complexidade Textual e Estrutura Textual.

As métricas de Complexidade Textual são referentes à ferramenta NILC-Metrix, sendo subdivididas em métricas de *Sofisticação Léxica*, *Complexidade Sintática*, *Dispositivos de Coesão* e *Tamanho e Extensão*.

A tabela 1 abaixo descreve as métricas utilizadas pela ferramenta Athos, divididas entre as categorias de Complexidade e Estrutura textual. A tabela apresenta o grupo a qual a métrica corresponde, o nome da métrica e um resumo descritivo sobre cada uma delas.

Tabela 1 - Apresentação das métricas da ferramenta Athos

	Grupo	Métrica	Resumo
Métricas de Complexidade Textual	Sofisticação Léxica	Diversidade léxica	Proporção de palavras sem repetições (types) em relação ao total de palavras com repetições (tokens).
		Frequência de palavras	Média das frequências das palavras de conteúdo do texto, via Banco de Português, Corpus Brasileiro e BrWac. Considerando palavras de conteúdo como substantivos, verbos, adjetivos e advérbios.
		Nível de concretude das palavras	Identificam-se as palavras de conteúdo do texto. Em seguida, lematizam-se essas palavras e procuram-se seus respectivos valores de concretude em um banco de dados específico para isso.
		Polissemia das palavras	Média de sentidos por palavra do texto, a quantidade de sentidos é o número de sentidos que uma palavra tem no dicionário TEP (Thesaurus Eletrônico do Português).
	Complexidade sintática	Palavras antes do verbo principal	Quantidade média de palavras antes dos verbos principais das orações principais das sentenças.
		Modificadores por sintagma nominal	Média dos tamanhos médios dos sintagmas nominais nas sentenças. Como há sintagmas nominais constituídos de outros sintagmas nominais, são computados apenas os de alto nível, ou seja, os mais próximos da raiz da árvore sintática.
	Dispositivos de coesão	Conectivos Lógicos	Proporção de conectivos causais em relação ao total de palavras do texto
		Sobreposição de palavras de conteúdo	Quantidade média de palavras de conteúdo que se repetem nos pares de sentenças adjacentes do texto.
	Tamanho e Extensão	Número de palavras por sentença	Quantidade de palavras por sentenças.
		Número de sentenças por parágrafo	Quantidade sentenças por parágrafo.

		Tamanho das Seções	Tamanho das seções de Introdução, Referencial Teórico e Metodologia do trabalho.
Métricas de Estrutura Textual	Coerência	Coe_Título-Objetivo	Análise da correlação entre título e objetivo.
		Coe_Referencial	Análise da correlação entre referencial teórico e palavras-chave do projeto.
		Coe_Elem-Palavras Chave	Análise da correlação entre os elementos estruturantes do projeto (título, objetivo e palavras-chave).
	Objetivo	Obj_Verbos	Análise da correlação dos verbos do objetivo.
		Obj_QuantAções	Análise da quantidade de ações do objetivo.
		Obj_AspecMetodol	Análise do objetivo e existência de aspectos metodológicos.
		Obj_TermosVagos	Análise da existência de palavras e pronomes vagos do objetivo.
	Metodologia	Met_Modalidade	Análise da modalidade da pesquisa.
		Met_ColetaDados	Análise dos procedimentos de coleta de dados.
		Met_Classificação	Análise da classificação do objetivo pesquisa.
		Met_Abordagem	Análise da classificação da abordagem da pesquisa.
		Met_PalavrasChave	Análise da relação entre metodologia e palavras-chave do projeto.

		Met_Sujeitos	Análise da descrição dos sujeitos da pesquisa.
		Met_Período	Análise do período de realização da pesquisa.

Fonte: do autor

Para fins desta pesquisa, a ferramenta Athos foi implementada e disponibilizada em servidor local. Está completamente operante e futuramente deverá ser disponibilizada em servidor online para uso aberto pela comunidade.

3.2 Etapa 2: Análise de métricas de complexidade e estrutura textual

Para a configuração das análises textuais realizadas pela ferramenta Athos, foram consideradas diversas métricas de análise textual para observar como se relacionam com avaliações feitas por professores e revisores. O primeiro estudo foi composto por redações avaliadas por meio dos critérios de avaliação do ENEM. O segundo estudo, por sua vez, abordou trabalhos de projetos de pesquisa desenvolvidos por alunos de graduação em disciplinas de metodologia da pesquisa.

Cada um dos estudos permitiu identificar as médias para diferentes métricas e avaliar a existência de correlações estatísticas entre o comportamento dessas médias e das notas atribuídas pelos avaliadores (professores e revisores, no caso das redações).

3.2.1 Metodologia para a Análise de Redações seguindo critérios do ENEM

O primeiro estudo desenvolvido analisou as redações do Banco de Redações do Portal UOL (<https://educacao.uol.com.br/bancoderedacoes/>), seguindo abordagem quantitativa. Foram coletadas 738 redações divididas entre 50 temas. As redações foram escritas e enviadas por internautas, sendo que a seleção e a avaliação das redações couberam a professores associados ao Banco de Redações. A correção traz a nota atribuída à redação, bem como a nota alcançada em cada competência, e algumas redações ainda apresentam um feedback separado por competência. Todas as redações possuíam comentários dos professores quanto aos pontos para a obtenção da nota atribuída.

O estudo das redações obtidas através do Banco de Redações do Portal UOL utilizou a ferramenta NILC-Metrix para a obtenção das métricas de análise textual desses trabalhos. Essa ferramenta já foi utilizada e validada em diferentes pesquisas com diferentes aplicações e finalidades, como por exemplo, na avaliação de textos lidos e ouvidos por crianças, destacando as diferenças entre as legendas de filmes e séries infantis do projeto Leg2Kids e no desenvolvimento de um novo preditor de complexidade textual para o corpus de textos originais e simplificados do projeto PorSimples (Leal, 2023).

A finalidade do estudo das redações une duas ideias. Primeiramente, a avaliação de textos, ou seja, análise e atribuição de nota para as redações empregando os parâmetros do ENEM. Em segundo lugar, a compreensão de como as métricas de complexidade textual se comportam para esse tipo de texto, observando também a possibilidade de predição das notas atribuídas originalmente pelos revisores.

A análise do estudo considerou tanto a nota final atribuída à redação, quanto a nota de cada competência. Aqui cabe esclarecer que a nota final da redação pode variar entre 0 e 1000 pontos e é composta pela soma das notas de cada uma das cinco competências, que podem variar de 0 a 200 pontos.

Após a coleta e processamento textual das redações, obtiveram-se os dados quantitativos correspondentes a esse corpus. Para analisar os resultados obtidos, observou-se a variação e correlação das métricas apresentadas na literatura com as notas atribuídas às redações. Essa análise foi feita através do teste estatístico ANOVA.

3.2.2 Resultados para a Análise de Redações seguindo critérios do ENEM

O teste estatístico ANOVA, executado no estudo das Redações, mostrou resultados divergentes da literatura, como os apresentados em Meira et. al. (2023) e McNamara e Graesser (2012). A tabela 2 abaixo apresenta a correlação de resultados ANOVA para as notas de cada competência e também para a nota final atribuída à redação.

Para os cálculos ANOVA no corpus de redações, as notas foram categorizadas da seguinte forma: Nota final da redação variando de 0 a 1000, sendo

divididas em 5 categorias de 200 pontos: categoria A de 1000 a 800 pontos; categoria B de 799 a 600 pontos; categoria C de 599 a 400 pontos; categoria D de 399 a 200 pontos; e categoria E de 199 a 0 pontos.

A categorização das notas das competências, que variam de 0 a 200 pontos, foi dividida em uma escala de 40 pontos, sendo categoria A de 200 pontos, categoria B de 160 pontos, categoria C 120 pontos, categoria D de 80 pontos, categoria E de 40 pontos e categoria F de 0 pontos.

Tabela 2 - Tabela de dados ANOVA Notas x Métrica

		P-Value					
		Nota Final	Comp 1	Comp 2	Comp 3	Comp 4	Comp 5
Sofisticação léxica	Diversidade léxica	= 0.0001	= 0.0001	= 0.0002	= 0.0007	= 0.0003	= 0.0003
	Frequência de palavras	= 0.21	= 0.857	= 0.702	= 0.0903	= 0.21	= 0.486
	Concretude das palavras	= 0.0033	= 0.0287	= 0.0004	= 0.0024	= 0.0025	= 0.0023
	Polissemia das palavras	= 0.0002	= 0.0001	= 0.0001	= 0.0008	= 0.0004	= 0,0005
Complexidade de Sintática	Palavras Antes do verbo principal	= 0.741	= 0.017	= 0.211	= 0.0582	= 0.0393	= 0.0595
	Modificadores por sintagma nominal	= 0.614	= 0.55	= 0.801	= 0.624	= 0.645	= 0.819
Dispositivos de Coesão	Conectivos lógicos	= 0.0454	= 0.122	= 0,369	= 0.117	= 0.28	= 0.187
	Sobreposição de palavras de conteúdo	= 0.0006	= 0.0005	= 0.0002	= 0,0002	= 0.0004	= 0.0001

Fonte: do autor

A tabela destaca que as métricas *Diversidade léxica*, *Concretude das palavras*, *Polissemia das Palavras*, *Conectivos Lógicos* e *Sobreposição de palavras de conteúdo* apresentaram correlação estatística com a nota final atribuída pelo avaliador. As análises por competência também revelam padrões interessantes,

ilustrando como diferentes aspectos da escrita são valorizados em diferentes dimensões da avaliação.

Notadamente, a categoria *Sofisticação Léxica* emergiu como altamente relevante, sugerindo que, apesar da métrica *frequência de palavras* não apresentar correlação estatística para todas as notas, coletivamente elas relacionam-se à avaliação final. Enquanto isso, a categoria Complexidade Sintática mostrou correlações limitadas, indicando uma potencial diferença no impacto dessas métricas em análises de redações em português comparado com estudos anteriores em inglês.

A categoria Dispositivos de Coesão mostrou-se robusta, estabelecendo correlações estatísticas significativas em quase todas as notas atribuídas, exceto para a métrica *Conectivos lógicos*. Dessa forma, a categoria mostra-se dividida, com uma métrica mostrando correlação para todas as notas e outra mostrando correlação apenas para a nota final. Esse resultado difere das observações de McNamara e Graesser (2012), que notaram uma correlação negativa entre o uso de dispositivos de coesão e as notas, possivelmente indicando que uma maior coesão pode diminuir a complexidade percebida do texto.

Ao analisar os resultados estatísticos obtidos pelas métricas em cada competência avaliativa, nota-se que a maioria das métricas apresentou relevância estatística na correlação com as notas atribuídas pelos avaliadores. Apesar das correlações encontradas, para uma análise adequada, é necessário considerar a que se refere cada uma das competências. As competências avaliadas nas redações do ENEM são expostas na cartilha do INEP, sendo divididas em cinco competências, cada uma com seis níveis de pontuação.

A competência 1 avalia se o participante demonstra domínio da norma culta da língua portuguesa, já a competência 2 trata da compreensão da proposta da redação e aplicação dos conceitos das várias áreas de conhecimento para desenvolver o tema dentro dos limites estruturais do texto dissertativo-argumentativo. A competência 3 por sua vez verifica se o participante consegue selecionar, relacionar, organizar e interpretar informações, fatos, opiniões e argumentos em defesa de um ponto de vista. Na competência 4 é avaliado se o participante demonstra conhecimento dos mecanismos linguísticos necessários para

a construção da argumentação e por fim, na competência 5 é avaliado se o participante elabora a proposta de intervenção para o problema abordado, respeitando os direitos humanos. (Brasil, 2023).

Ao se analisar o objetivo de cada competência, nota-se que, como no caso da competência 5, as ferramentas de PLN podem não conseguir capturar e analisar quanto o texto respeitou os direitos humanos. Para uma análise mais aprofundada dessa questão seria necessário uso de ferramentas de IA ou sistemas mais robustos de aprendizagem de máquina.

Assim, mesmo que tenham sido encontradas correlações entre as métricas e as notas atribuídas para a competência 5, talvez essa correlação não esteja ligada diretamente ao objetivo avaliado pela competência.

Por sua vez, as competências 1 e 4 objetivam avaliar diretamente os mecanismos linguísticos e o domínio da linguagem escrita por parte do candidato, sendo que os achados mais importantes do estudo são as correlações da nota final e das competências 1 e 4.

Esses resultados incentivam a continuação da investigação sobre como as métricas textuais interagem com as avaliações em redações, especialmente em língua portuguesa. A necessidade de expandir estudos em gêneros textuais diversos com alunos brasileiros é evidente, prometendo enriquecer ainda mais nossa compreensão da escrita e suas nuances.

Comparando os resultados estatísticos do estudo com os trabalhos de McNamara (2011), McNamara e Graesser (2012) e Meira et. al. (2023), é possível analisar cada categoria de métricas separadamente. A categoria *Sofisticação Léxica* foi a que mais demonstrou significância estatística no corpus de redações. Mesmo que a métrica *Frequência de palavras* não tenha apresentado consenso estatístico para todas as notas, pode-se entender que o conjunto da categoria tem relevância na variação da nota.

A categoria *Complexidade Sintática*, por sua vez, quase não mostrou relevância, tendo a métrica *Palavras antes do verbo principal* correlação com algumas notas de competências. Isso difere do reportado em trabalhos com análises em língua inglesa, uma vez que não verifica-se que a variação das métricas dessa

categoria influenciam a nota das competências e conseqüentemente a nota final da redação.

Para a categoria *Dispositivos de Coesão*, as métricas apresentaram significância estatística para a maioria das notas atribuídas, ficando sem correlação apenas a métrica *Conectivos lógicos* nas avaliações das competências. McNamara e Graesser (2012) indicaram uma correlação negativa entre o maior uso desses dispositivos e as notas dos trabalhos, sugerindo que, em certos casos, um aumento na coesão pode prever uma diminuição na complexidade do texto, o que pode impactar na avaliação da métrica.

Apesar de as métricas de *Sofisticação Léxica* e *Dispositivos de Coesão* terem apresentado relevância estatística significativa na predição das notas das competências, os resultados não são conclusivos. Deste modo, compreende-se necessário investigar mais detalhadamente como as métricas textuais podem estar relacionadas ao modo como os professores e revisores atribuem notas para as competências consideradas na avaliação de redações no modelo definido pelo ENEM.

Um gargalo da área de análise textual correlacionada com avaliações de especialistas encontra-se no fato de que a maioria dos trabalhos existentes abordam textos em língua inglesa, aliando-se a necessidade de estudos em língua portuguesa com alunos brasileiros.

3.2.3 Metodologia para a Análise dos Projetos de Pesquisa de Graduação

O segundo estudo realizado para avaliação do uso de métricas de análise textual pela ferramenta Athos envolveu projetos de pesquisa de graduação de disciplinas de Pesquisa em Educação oferecidas pela Faculdade de Educação da UFRGS para diferentes cursos de licenciatura. Também obtiveram-se dados sobre a estrutura textual a partir de componentes oriundos da ferramenta MTA. De posse desses dados, deu-se início ao estudo que também seguiu abordagem quantitativa.

Foram coletados 206 trabalhos de 23 turmas semestrais entre os anos de 2019 e 2023. Os trabalhos continham avaliações atribuídas pelo professor, com os conceitos A (ótimo), B (bom) e C (regular). Não havia trabalhos com conceitos D (insuficiente) na amostra.

Com esse estudo foi possível efetuar a correlação de ocorrência das métricas textuais com a nota atribuída pelo professor da disciplina, a importância dessa ação se resume na comparação e corroboração dos resultados para com a literatura, permitindo que as métricas escolhidas estejam devidamente embasadas.

3.2.4 Resultados para a Análise dos Projetos de Pesquisa de Graduação

As métricas avaliadas neste estudo seguiram a padronização da literatura, também disponíveis na tabela 1, a qual também foi aplicada no outro estudo desta pesquisa. Na tabela 3 são apresentados os resultados obtidos durante o estudo dos projetos de pesquisa para as métricas de Complexidade Textual. Neste estudo foram efetuados cálculos estatísticos com objetivo de encontrar uma correlação entre a nota atribuída pelo professor e as métricas extraídas pelas ferramentas NILC-Metrix.

Tabela 3 - Quadro de resultados para o corpus Projetos de Pesquisa

Métricas	Geral		Conceito A		Conceito B		Conceito C	
	Média	DP	Média	DP	Média	DP	Média	DP
Diversidade léxica	1,0184	0,3412	1,0679	0,3593	0,8756	0,3495	0,7994	0,2299
Frequência de palavras	62193	41262	639341	342825	583101	659233	482072	198519
Concretude das palavras	6,8367	4,3544	7,6121	4,8486	4,8042	2,1064	4,3561	1,2123
Polissemia das palavras	7,1265	4,2936	7,8604	4,7880	5,1763	2,1735	4,7352	1,4087
Palavras antes do verbo principal	10,412	7,1932	11,517	7,9535	7,5796	4,4333	6,6524	2,2121
Modificadores por sintagma nominal	10,901	7,5526	12,239	8,3691	7,6144	3,7672	6,2996	2,2827
Conectivos Lógicos	0,0683	0,0268	0,0705	0,0226	0,0602	0,0263	0,0590	0,0284
Sobreposição de palavras de conteúdo	1,8510	2,1187	2,0134	2,4302	1,4542	1,0095	1,2091	0,6229

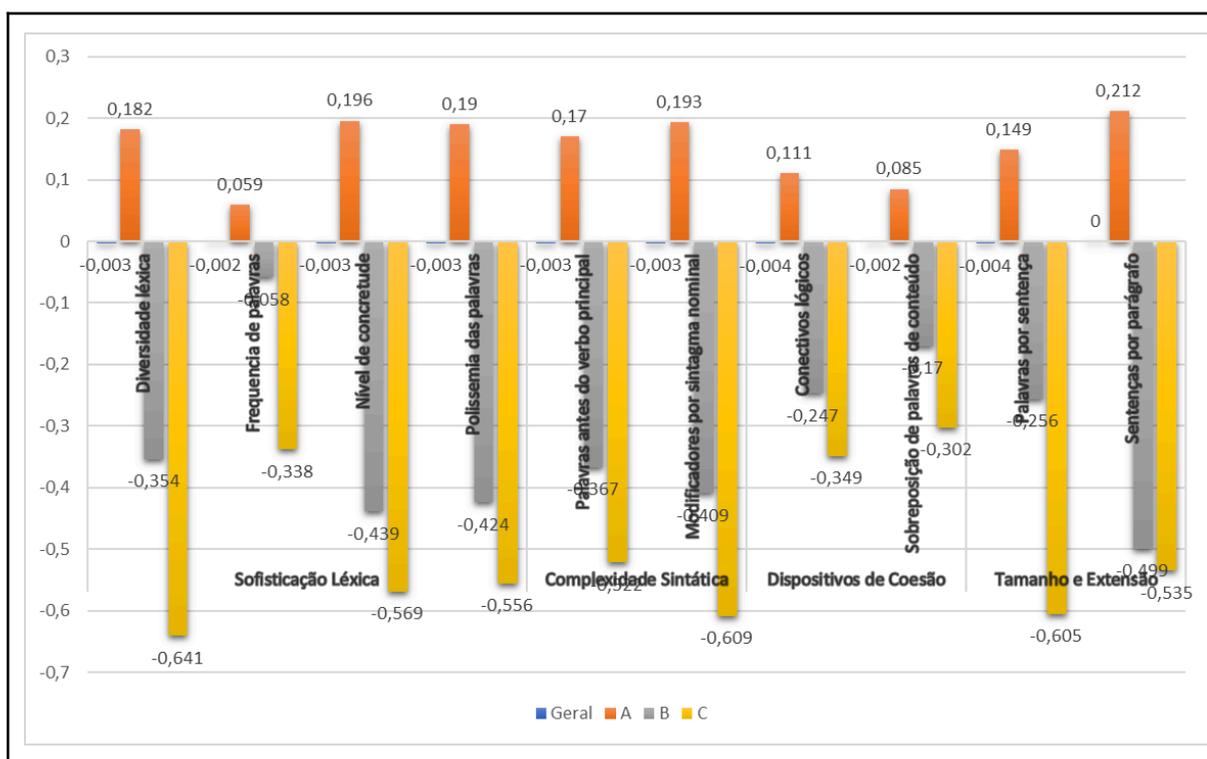
Fonte: do autor

A leitura da tabela 3 demanda lembrar que os valores atribuídos às colunas Desvio Padrão (DP) são referentes às médias que os acompanham, ou seja, para a métrica *Concretude das Palavras*, tem-se um desvio padrão a cada 4,35 pontos, para mais ou para menos. Essas informações foram importantes para a parametrização dos feedbacks gerados pelo sistema Athos.

Na tabela 3, é mostrada uma redução nas médias das métricas em relação aos conceitos A, B e C. Essa observação é consistente com os estudos de McNamara e Graesser (2012), exceto no que diz respeito às métricas *Frequência de Palavras*, *Conectivos Lógicos Positivos* e *Sobreposição de Palavras*, onde eles notaram uma correlação inversa positiva com as avaliações dos trabalhos. Isso indica a possibilidade de tendências divergentes no impacto desses aspectos textuais nas avaliações dos trabalhos.

Para facilitar a visualização das disposições das métricas quanto à média e aos respectivos conceitos atribuídos pelo professor, apresenta-se na figura 14 a disposição das métricas através do valor do z-score médio. Cada conceito foi comparado com a média geral dos trabalhos.

Figura 14 - Disposição de z-scores por conceito



Fonte: do autor

Complementando a leitura da tabela 3, mostram-se os valores de média e de desvios padrão dos trabalhos dentro das suas respectivas avaliações. A figura 14 apresenta o comparativo das médias de pontuação z-score dos trabalhos de cada categoria de avaliação (A, B e C) para com a média geral dos trabalhos.

A leitura complementar da tabela 3 juntamente com a figura 14 salienta as diferenças encontradas na avaliação das métricas dos trabalhos divididos por conceitos. Os trabalhos que obtiveram conceito “A” obtiveram média de valor z-score positivo para todas as métricas. Os trabalhos de conceito “C”, por sua vez, obtiveram média de valor z-score negativo, sendo inferiores aos valores médios obtidos por trabalhos de conceito “B”.

Com a leitura da figura 14 conclui-se que uma redução no valor obtido para as métricas selecionadas aponta para a direção de avaliações piores por parte dos professores, ou seja, trabalhos avaliados com o conceito A, obtiveram valor z-score predominantemente positivos à média. Já trabalhos classificados com conceito B obtiveram pontuação z-score entre -0,25 e -0,44. Por sua vez, os trabalhos avaliados como conceito C obtiveram para a maioria das métricas de complexidade textual, valores z-score inferiores a -0,5.

Esses resultados mostram uma capacidade das métricas na avaliação e predição da nota atribuída pelo professor para textos do gênero projetos de pesquisa, porém para uma melhor compreensão da participação individual das métricas na variação dos conceitos, buscou-se explorar a sua relação estatística.

Para explorar a existência de relações significativas entre as métricas mencionadas e as avaliações dos trabalhos do corpus, foi conduzida uma análise de variância (ANOVA) entre as métricas dos trabalhos e os grupos formados pelas categorias de conceito (A, B e C). Essa análise buscou verificar se as diferenças observadas eram estatisticamente relevantes e se as variáveis independentes se correlacionavam com a variável categórica 'conceito'. Os achados desta análise são apresentados na tabela 4. Os resultados da análise foram publicados em Meira et. al. (2023), e maiores detalhes sobre as correlações encontradas são descritos no artigo.

Tabela 4 - Tabela de correlações para notas e métricas

	Média	Desvio Padrão	F-Value	P-Value
Diversidade léxica	1,0184	0,3412	9,993	= 0,001
Frequência de palavras	62193	41262	1,5172	= 0,221

Concretude das palavras	6,8367	4,3544	11,0320	= 0,001
Polissemia das palavras	7,1265	4,2936	10,3343	= 0,001
Palavras antes do verbo principal	10,4127	7,1932	8,1911	= 0,001
Modificadores por sintagma nominal	10,9012	7,5526	10,9046	= 0,001
Conectivos lógicos	0,06838	0,0268	3,5378	= 0,032
Sobreposição de palavras de conteúdo	1,8510	2,1187	2,0977	= 0,117

Fonte: do autor

Os resultados apresentados na tabela 4 indicam correlações estatísticas significativas em várias métricas analisadas, com valores de p-value menores que 0,01, indicando alta significância estatística. Entretanto, três métricas específicas – *Frequência de Palavras*, *Conectivos Lógicos Positivos* e *Sobreposição de Palavras de Conteúdo* – não apresentaram significância estatística, divergindo dos achados típicos em análises de textos em inglês, como apresentado em Graesser e McNamara (2011). Separadamente, a *Sofisticação Léxica* mostrou uma clara associação entre altas avaliações e maior complexidade lexical, embora a métrica *Frequência de Palavras* não tenha mostrado significância. Isso sugere que, apesar da ausência de correlação com a frequência das palavras, textos com maior complexidade lexical tendem a ser mais bem avaliados, mesmo que mais difíceis de entender.

Em relação à *Complexidade Sintática*, os resultados estão alinhados com os estudos anteriores (McNamara e Graessner, 2012), indicando que estruturas sintáticas mais complexas estão positivamente correlacionadas com avaliações mais altas. Por outro lado, na categoria *Dispositivos de Coesão*, não foram encontradas correlações significativas entre o uso desses dispositivos e as avaliações dos trabalhos. Isso contrasta com as tendências reportadas por McNamara e Graessner (2012), sugerindo que as diferenças podem ser devido ao tipo de texto ou às características específicas do português comparadas ao inglês.

Ao efetuar comparação destes dados de Projetos de Pesquisa com os dados de Redações, notou-se que a diferença dos estilos de textuais e objetivos da escrita pode influenciar positiva ou negativamente os resultados obtidos. Apesar de correlações estatística significantes terem sido encontradas no estudo de Redações,

os resultados não foram conclusivos ao relacionar a nota da redação com os valores obtidos através das métricas, diferentemente dos resultados apresentados neste estudo.

Assim, a variação da correlação apresentada pelos testes ANOVA dos estudos de projetos de pesquisa e de redações é atribuída à diferença de estilo dos textos existentes em cada corpus. Nos projetos de pesquisa, tem-se a média de 1968,2 palavras, 99,8 sentenças e 52,7 parágrafos, já para as redações, as médias obtidas, respectivamente, são 257,8, 9,6 e 4,2. Estas diferenças básicas de tamanho textual já demonstram diferenças na forma e no objetivo da escrita, sendo que isto tende a limitar o uso de palavras categorizadas dentro das métricas avaliadas.

Outro ponto de possível motivador para as diferenças estatísticas está na metodologia de avaliação dos trabalhos dos estudos, uma vez que as redações possuem características onde o contexto e a evolução do aluno não são julgados pelo avaliador.

Com base nestes resultados apresentados, entende-se que o aprofundamento nos estudos com variações de gêneros textuais são necessários para o refinamento das métricas utilizadas, assim como a busca de correlação entre redações e outras métricas não utilizadas neste estudo. A implementação de diferentes técnicas de PLN também podem apresentar variação nos resultados.

Para o estudo de projetos de pesquisa, ainda foram calculadas as correlações existentes entre as notas atribuídas pelo professor da disciplina e as métricas de Estrutura Textual obtidas através do Plugin MTA.

Os resultados obtidos pelo estudo de Projetos de Pesquisa, apresentados na tabela 5, demonstram a taxa de ocorrência percentual de cada métrica relacionada à Estrutura Textual, para este corpus. A tabela está separada em colunas correspondentes às métricas de cada categoria, detalhadas na tabela 1, e em linhas, representando Pos para as ocorrências positivas e Neg para as ocorrências negativas das respectivas métricas.

Tabela 5 - Ocorrência das Métricas de Estrutura Textual

	Coerência			Objetivo			
	Títu-Objetiv	Referencial	Elem-PalavrCh	Verbos	QuantAções	AspecMeto	TermosVago
Pos	31,72	57,36	45,51	84,14	0	0	0
Neg	22,76	42,64	0	14,48	10,34	3,45	0,69

	Metodologia						
	Modalidade	ColetaDados	Classificação	Abordagem	PalavrChav	Sujeitos	Período
Pos	84,14	77,93	35,17	17,93	32,41	88,97	82,76
Neg	15,86	22,07	64,83	82,07	66,21	11,03	17,24

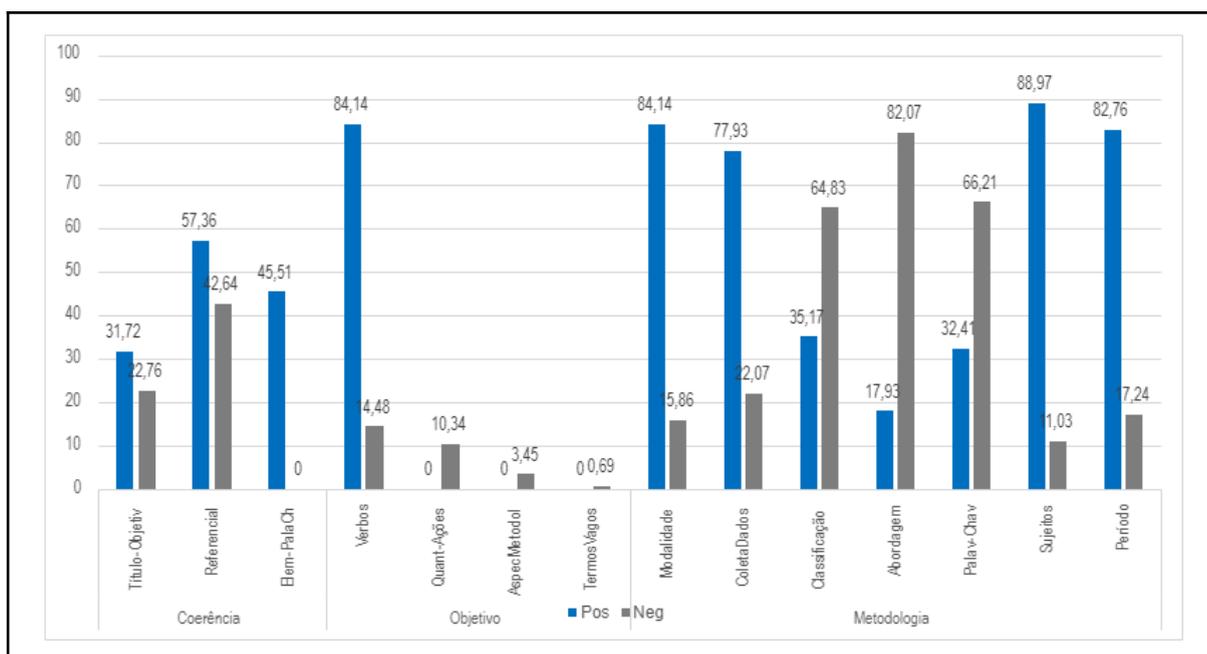
Fonte: do autor.

Na tabela 5, não se verifica a ocorrência de retornos negativos para a métrica “*Elem-PalavrCh*” do quesito coerência, a lembrar, esta métrica é correlacionada à métrica “Título-Objetivo” do mesmo quesito, sendo que possui apenas retorno positivo. Desta forma, quando a métrica “*Título-Objetivo*” tem resultado positivo e a métrica “*Elem-PalaChav*” tem resultado negativo, apenas o feedback positivo é apresentado, evidenciando apenas a correlação entre título e objetivo. Porém, quando ambas as métricas têm resultado positivo, apenas o texto relativo à métrica “*Elem-PalaChav*” é apresentado, já que ele condensa o feedback de ambas as métricas. Esta configuração tem o objetivo de resumir o feedback final apresentado.

Hattie e Timperley (2007) comentam que ao oferecer feedbacks construtivos, é mais importante que este feedback seja focado na qualidade e relevância do que na repetição e extensão. Assim, o excesso de textos no feedback pode dificultar a compreensão dos pontos a serem melhorados.

Para facilitar a leitura da tabela 5, construiu-se a figura 15. Esta figura contém a taxa percentual de ocorrência de cada Métrica de Estrutura Textual para o corpus em questão.

Figura 15 - Gráfico de ocorrência das métricas de ET - Projetos



Fonte: do autor.

A figura 15 ilustra as ocorrências de retornos positivos e negativos através dos feedbacks gerados.

As métricas do quesito Objetivo, com exceção da métrica “*Verbos*”, não geram retornos positivos aos alunos, ou seja, não apresentam feedbacks parabenizando o aluno por escrever seu objetivo com apenas uma ação, por não mencionar aspectos metodológicos ou por não inserir termos vagos, por exemplo.

Para buscar a correlação das métricas de Estrutura Textual com as notas atribuídas pelo professor da disciplina, efetuou-se o cálculo da correlação de Pearson. Este teste permite identificar a força e a direção entre os dados analisados.

Para a execução do teste foram utilizadas como variáveis a nota atribuída pelo professor e a taxa de ocorrência das métricas. Como forma de entender melhor os dados, foram acrescentadas as compilações dos grupos de métricas ao cálculo. A tabela 6 apresenta os resultados.

Tabela 6 - Correlação de Pearson para métricas de Estrutura Textual

	Coeficiente de Correlação	p-value
Métricas Coerência Agregadas	0.0860	0.3324

Métricas Objetivo Agregadas	0.5074	0.0959
Métricas Metodologia Agregadas	0.7942	0.0471
Título-Objetivo	- 0.0264	0.7656
Referencial-PalavrasChave	- 0.0646	0.4669
Elementos-PalavrasChave	- 0.0409	0.6446
Objetivo-Verbos	0.0216	0.8073
Objetivo-QuantidadeAções	0.1087	0.2199
Objetivo-AspectosMetodologicos	0.2378	0.0670
Objetivo-TermosVagos	0.1526	0.5534
Metodologia-Modalidade	0.3562	0.0918
Metodologia-ColetaDados	0.2875	0.0795
Metodologia-Classificação	0.4250	0.0485
Metodologia-Abordagem	0.1672	0.2844
Metodologia-PalavrasChave	0.2135	0.3502
Metodologia-Sujeitos	0.0532	0.8752
Metodologia-Período	0.0175	0,6725

Fonte: do autor.

Lê-se na tabela 6 que as métricas de maneira individual não apresentaram indicativos estatísticos de rejeição da hipótese nula, além de apresentar uma baixa correlação entre a ocorrência da métrica e a nota atribuída pelo professor. Os resultados de correlação de Pearson mais próximos de 0 para a maioria das métricas de maneira individual indicam pouca força destas na correlação para com as notas atribuídas pelo professor. Este entendimento é reforçado pelo valor p maior que 0.05, indicando baixa significância estatística na rejeição da hipótese nula, ou seja, não é possível indicar que uma métrica de maneira isolada teve influência significativa para prever a nota atribuída pelo professor.

Este fato está alinhado aos achados de McNamara e Graesser (2012) e Meira et. al. (2023), apontando para o fato de que as categorias das métricas obtiveram melhores resultados na predição das notas do que as métricas de maneira

individual, como visto na tabela 6, onde as métricas de maneira agrupada obtiveram melhores resultados.

Outro fator consiste na função das métricas de Estrutura Textual. Cada uma delas avalia um ponto específico de um projeto de pesquisa. Desta forma, não era esperado que apenas uma métrica predissesse a nota atribuída pelo professor. Os resultados apresentados na tabela 6 reforçam este entendimento, sendo possível afirmar que estas métricas não têm peso determinante de maneira individual.

Considerando todas as métricas de metodologia para cada trabalho como uma métrica agrupada, a correlação entre agrupamento e nota do professor apresentou resultados mais fortes na correlação de Pearson, obtendo valor de 0.79 para a força da correlação, o que indica uma correlação forte. No trabalho de Mukaka (2012), este agrupamento ainda apresentou relevância estatística significativa, obtendo grau superior a 95%, indicado através do valor p.

Outra métrica agrupada que apresentou alguma correlação entre agrupamento e nota, porém desta vez com força moderada segundo a escala apresentada em Mukaka (2012), foi o agrupamento de métricas Objetivo. Com valor de correlação de 0.5074 e valor p em 0.0959 é possível entender que há alguma correlação entre o agrupamento e a nota atribuída pelo professor, porém não é possível afirmar que a hipótese nula está descartada, uma vez que o grau de significância estatística ficou abaixo de 95%.

Já para o agrupamento de métricas de Coerência não se vislumbrou nenhuma correlação entre o agrupamento e as notas atribuídas, uma vez que os valores obtidos são categorizados como sem correlação por Mukaka (2012). Assim, entende-se necessário continuar os estudos sobre o uso destas métricas com maior número de trabalhos, a fim de confirmar ou refutar a correlação encontrada.

3.3 Etapa 3: Estudo Final para Avaliação do Emprego da ferramenta Athos

A terceira etapa da pesquisa foi iniciada deu-se com a utilização da ferramenta Athos. As seções a seguir detalham aspectos metodológicos do estudo e em seguida os resultados alcançados.

3.3.1 Metodologia do Estudo Final para Avaliação do Emprego da ferramenta

Esta fase do estudo foi desenvolvida na modalidade de estudo de casos múltiplos. Essa é uma metodologia de pesquisa qualitativa que envolve o estudo de vários casos para compreender um fenômeno complexo, permitindo aos pesquisadores examinar padrões, semelhanças e diferenças entre os casos. (Yin, 2014)

O contexto de uso e avaliação da ferramenta consistiu em classes de disciplinas de metodologia da pesquisa em educação, disciplinas de cursos de nível superior em grau de licenciatura de diferentes áreas acadêmicas. Essas turmas contam com média de 25 alunos com entrada semestral. Três professores dessas disciplinas participaram dos estudos. Todos eram experientes em ministrar disciplinas correlatas e em avaliar trabalhos científicos. Um total de 5 turmas foi envolvido na pesquisa.

As atividades que foram realizadas com os alunos consistiram de atividades usuais das disciplinas, que têm como foco central a escrita de um projeto de pesquisa no formato de uma proposta de trabalho de conclusão de curso. Durante a disciplina, os alunos habitualmente trabalham com diferentes atividades em que vão desenvolvendo diferentes partes de seus projetos, a saber:

- I. Introdução
 - A. Tema da pesquisa
 - B. Questão de pesquisa
 - C. Objetivos
 - D. Justificativa
- II. Revisão de literatura
 - A. Autores/teorias que dão suporte teórico ao estudo
 - B. Trabalhos relacionados à proposta
- III. Metodologia

Caracterização dos procedimentos metodológicos com relação à:

- A. natureza da pesquisa (descritiva, explicativa, exploratória)
- B. abordagem (quantitativa, qualitativa)
- C. modalidade (estudo de caso, pesquisa experimental, pesquisa etnográfica, pesquisa-ação, pesquisa de levantamento, ...)

Descrição dos procedimentos metodológicos com relação à:

- D. coleta de dados (entrevistas, questionários, grupos focais, logs de utilização de sistemas, ...)
- E. análise de dados (análise estatística, análise de conteúdo, análise de discurso)

As atividades monitoradas pela pesquisa foram desenvolvidas pelo professor seguindo o cronograma da disciplina. Durante as entregas das atividades ao longo do semestre, o professor fez uso da ferramenta Athos para dar retorno aos alunos sobre seus trabalhos. Foram previstas as seguintes fases de avaliação:

- **5a semana de aula:**
 - alunos entregaram versão do trabalho que apresenta o título, discorreram sobre o tema, a questão de pesquisa, o objetivo geral; listaram ainda os assuntos a serem tratados no referencial teórico
 - professor avaliou atividade e retornou aos alunos o trabalho comentado
- **10a semana de aula:**
 - alunos entregaram nova versão incluindo agora a revisão de literatura que já tinha os temas definidos
 - professor avaliou atividade e retornou aos alunos comentários sobre todo o texto, com foco na revisão de literatura e nos comentários enviados na avaliação anterior
- **16a semana de aula:**
 - alunos entregaram nova versão incluindo agora a seção de metodologia da pesquisa
 - professor avaliou atividade e retornou aos alunos comentários sobre todo o trabalho, focando principalmente na seção de metodologia e nos comentários enviados na avaliação anterior
- **17a semana de aula:**
 - alunos entregaram a versão final e revisada do trabalho
 - professor revisou novamente o trabalho, focando em aspectos que haviam sido comentados na versão anterior

Em cada trabalho avaliado, para todas as 4 fases de avaliação, foram coletados dados nos logs do sistema sobre como os professores utilizaram os feedbacks produzidos automaticamente pela ferramenta:

- feedback utilizado na íntegra, sem nenhuma alteração
- feedback alterado em até 15%
- feedback alterado em até 50%
- feedback alterado em mais de 50%
- feedback desconsiderado, registrado como inadequado

Durante o uso da ferramenta para avaliação dos trabalhos, os professores também dispuseram de um campo de texto para poderem comentar os feedbacks recebidos pela ferramenta e foram incentivados a registrar seus comentários. Esses comentários foram examinados por meio de análise de conteúdo, seguindo diretrizes estabelecidas por Saldaña (2013) e descritas mais adiante no texto.

No final do semestre, foi realizada uma análise dos logs de uso da ferramenta pelos professores, a maneira como os feedbacks foram utilizados bem como os comentários feitos pelos professores para os feedbacks. A partir desses dados, foi elaborada uma entrevista semi-estruturada para melhor conhecer a percepção dos professores sobre o apoio da ferramenta na realização das avaliações dos projetos. De modo geral, buscou-se compreender se a ferramenta agilizou e qualificou o trabalho de avaliação, segundo a visão dos professores. Foram também localizados episódios de uso da ferramenta em que o professor aceitou os feedbacks na sua íntegra ou fez pequenas modificações, bem como aqueles em que o professor fez muitas modificações e/ou descartou o texto elaborado pela ferramenta completamente. Os professores foram então solicitados a comentar mais detalhadamente esses episódios. Além desses questionamentos, as seguintes perguntas fizeram parte do roteiro de entrevista:

- Em que medida a ferramenta facilitou ou agilizou o processo de fornecer feedbacks aos alunos?
- Quais recursos ou funcionalidades específicas da ferramenta foram mais úteis para você ao avaliar os projetos dos alunos?
- Houve alguma dificuldade ou limitação encontrada ao utilizar a ferramenta? Se sim, como você lidou com essas questões?

- Comparado ao método tradicional de fornecer feedbacks, você considera que a ferramenta melhorou a qualidade dos feedbacks que você pôde oferecer aos alunos? Por quê?
- A ferramenta ajudou a padronizar ou tornar mais consistente o processo de feedback em relação a diferentes projetos e alunos?
- Como você percebeu a mudança na receptividade dos alunos aos feedbacks encaminhados neste semestre?
- Você acredita que a ferramenta teve algum impacto na motivação ou engajamento dos alunos em relação aos projetos? Por quê?
- Como você tratou os gráficos comparativos apresentados? Eles foram úteis na avaliação dos trabalhos?
- Com base na sua experiência, você tem alguma sugestão de melhoria ou recursos adicionais que poderiam ser incluídos na ferramenta para torná-la ainda mais eficaz?

Após a realização e transcrição das entrevistas com os professores, essas foram analisadas através da técnica de análise de conteúdo com o processo de codificação descrito em Saldaña (2013). Esse processo conta com a criação de códigos identificadores que transcrevem a ideia do texto analisado, orientando assim a identificação e categorização do conteúdo. O autor trata um código na pesquisa qualitativa como sendo uma palavra ou frase curta, que confere um atributo somativo ou evocativo para uma porção de dados. A etapa de codificação contou com dois codificadores e no final dessa fase, foi efetuada a verificação de inconsistências dos dados entre eles. Chegando-se a um consenso, iniciou-se uma nova fase em que os códigos eram identificados nas entrevistas, tarefa novamente a ser realizada pelos dois codificadores, com um período de conciliação ao final do processo. Por fim, a ocorrência de cada código foi contabilizada nas falas dos professores, buscando-se identificar frequências e tendências. Trechos das falas foram escolhidos para ilustrar e discutir a ocorrência dos códigos, buscando-se identificar de que modo a ferramenta proposta auxiliou na qualificação do processo de mediação e avaliação dos projetos.

Em resumo, os métodos de coleta de dados empregados foram:

- Construção de corpora de dados balizadores das métricas de Complexidade Textual e de Estrutura Textual para as diferentes áreas.
- Qualificação feita pelo professor para o feedback produzido automaticamente pela ferramenta (usado na íntegra, alterado em até 15%, 50%, mais de 50%, desconsiderado), para cada trabalho avaliado
- Coleta e análise dos logs dos professores para organização das entrevistas semi-estruturadas identificando episódios em que feedbacks foram usados sem muitas alterações, e também aqueles em que foram muito alterados ou desconsiderados
- Entrevista semi-estruturada com professores para compreender de que modo a ferramenta auxiliou na qualificação do processo de mediação e avaliação dos projetos.

3.4.2. Resultados do Estudo Final para Avaliação do Emprego da ferramenta

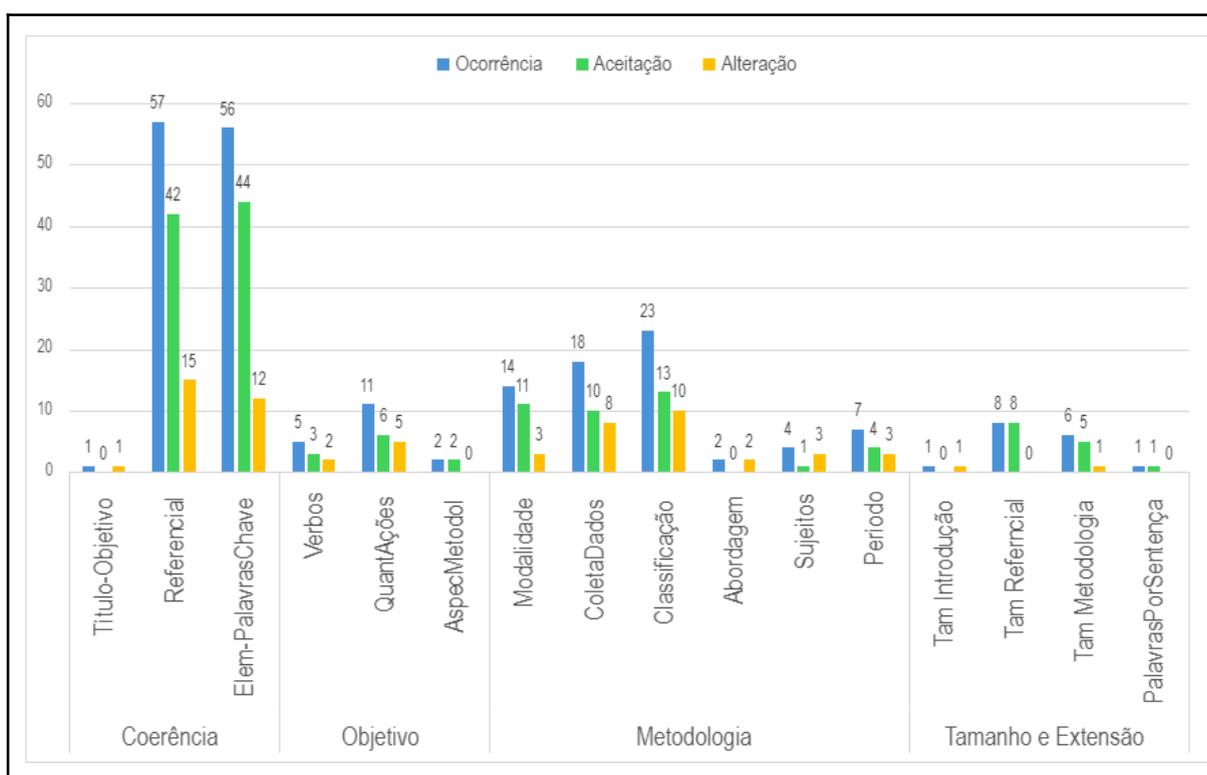
A primeira parte do estudo final para avaliação do emprego da ferramenta Athos apoiou-se no campo da pesquisa quantitativa. Os dados foram coletados após o uso da ferramenta Athos pelos três professores participantes da pesquisa. Esses dados se concentram em torno dos feedbacks gerados pelas análises da ferramenta para os trabalhos e na aceitação, alteração ou rejeição desses feedbacks por parte dos professores.

Para o controle de geração dos feedbacks, foram utilizados os dados obtidos com os estudos prévios realizados. Esses estudos permitiram e direcionaram a construção dos parâmetros de geração de feedback personalizado para os trabalhos submetidos, podendo a ferramenta Athos ser utilizada tanto por orientadores quanto por professores de disciplinas de metodologia científica ou da pesquisa.

Os feedbacks foram gerados baseando-se no desvio padrão de ocorrência das métricas dentro do estudo dos projetos de pesquisa. Para as métricas de Complexidade Textual, foram definidos parâmetros de variação do desvio padrão para cada métrica componente das categorias. A exemplificar, a métrica *Frequência de Palavras* da categoria Sofisticação Léxica somente dispara um feedback ao usuário quando o valor z-score da métrica variar mais do que 1.5, tanto para mais quanto para menos da média do corpus ou área, ou seja, quando o valor obtido pelo trabalho for um desvio padrão e meio acima ou abaixo da média.

Como forma de responder à questão norteadora desta pesquisa, uma das variáveis de coleta elencada resumiu-se à aceitação ou não dos feedbacks gerados por parte dos professores participantes. Assim, os dados quantitativos coletados através da ferramenta mostraram o percentual de aceitação dos feedbacks. A figura 16 abaixo apresenta a disposição de feedbacks gerados pela ferramenta, divididos entre a ocorrência dos feedbacks, o número de feedbacks aceitos e o número de alterações ou remoções efetuadas pelo professor, no feedback antes desse ser enviado ao aluno.

Figura 16 - Disposição dos feedbacks por métricas



Fonte: do autor.

Na figura 16 encontra-se a frequência de ocorrência de cada métrica disparadora de feedback, representado por barras azuis, bem como a aceitação dos feedbacks gerados, representado através das barras verdes, e o número de alterações nos feedbacks, representados pelas barras amarelas. O sistema contou com 57 trabalhos avaliados por três professores participantes, sendo o primeiro professor, avaliador de 28 trabalhos, o segundo professor com 23 avaliações e o terceiro com 6 trabalhos avaliados. Nota-se no gráfico que, assim como no estudo prévio de Projetos de Pesquisa, a métrica Coerência “Referencial” obteve 100% de

ocorrência. Essa taxa de ocorrência deve-se ao fato de que essa métrica possui tanto retorno positivo quanto negativo, já a métrica coerência “*Elem-PalavrasChave*”, por sua vez, deve ter sua taxa de ocorrência somada com a métrica Coerência “*Título-Objetivo*”, uma vez que essas métricas são correspondentes, ou seja, onde a métrica “*Título-Objetivo*” apresentar feedback corretivo, a métrica “*Elem-PalavrasChave*” não será disparada. Dessa forma, tem-se também 100% de ocorrência, sendo 98% para a Coerência “*Elem-PalavrasChave*” e 2% para a Coerência “*Título-Objetivo*”. As métricas apresentadas na tabela 1 e suprimidas na figura 16 não tiveram ocorrências registradas dentre os trabalhos avaliados pelos professores participantes.

Além das taxas de ocorrência das métricas para os professores participantes, a figura 16 mostra a validação dessa ocorrência por parte do professor. Nota-se nas barras verdes a aceitação do feedback gerado para cada métrica. Nesse caso, foram disparadas 216 vezes as métricas geradoras de feedback, dessas o professor validou 150 ocorrências, consistindo em um percentual de aceitação de 69,44%, restando 66 feedbacks alterados, removidos ou adicionados pelos professores, representados pelas barras amarelas.

A taxa de ocorrência de cada métrica é detalhada na tabela 7 abaixo, esta tabela é complementar à figura 16.

Tabela 7 - Taxa de aceitação dos feedbacks

	Ocorrência	Aceitação	Alteração	Percentual Aceitação
Título-Objetivo	1	0	1	0,00%
Referencial	57	42	15	73,68%
Elem-PalavrasChave	56	44	12	78,57%
Verbos	5	3	2	60,00%
QuantAções	11	6	5	54,55%
AspecMetodol	2	2	0	100,00%
Modalidade	14	11	3	78,57%
ColetaDados	18	10	8	55,56%
Classificação	23	13	10	56,52%
Abordagem	2	0	2	0,00%
Sujeitos	4	1	3	25,00%
Período	7	4	3	57,14%
Tam Introdução	1	0	1	0,00%
Tam Referencial	8	8	0	100,00%
Tam Metodologia	6	5	1	83,33%
PalavrasPorSentença	1	1	0	100,00%
Recusou na integra	2			3,51%
Aceitou na integra	22			38,60%
Taxa aceitação	216	150	66	69,44%

Fonte: do autor

A tabela 7 representa a taxa de aceitação dos feedbacks gerados para os trabalhos dos alunos avaliados pelos professores participantes da ferramenta. A descrição de cada métrica encontra-se na tabela 1. Para lembrar, a métrica Coerência “*Título-Objetivo*” analisa a correlação entre título e objetivo, a métrica Coerência “*Referencial*” avalia a relação entre as palavras-chave do trabalho com sua ocorrência e relação no referencial teórico, a métrica Coerência “*Elem-PalavrasChave*” considera a relação entre título, objetivo e palavras-chave, a métrica Objetivo “*Verbos*” analisa a correlação dos verbos do objetivo, a métrica

Objetivo “*QuantAções*” a quantidade de ações do objetivo, a métrica Objetivo “*AspecMetodoI*” considera a existência de aspectos metodológicos no objetivo, a métrica Metodologia “*Modalidade*” analisa a modalidade da pesquisa, a métrica Metodologia “*ColetaDados*” refere-se a coleta de dados, a métrica Metodologia “*Classificação*” aborda a classificação do objetivo da pesquisa, a métrica Metodologia “*Abordagem*” analisa a abordagem da pesquisa, a métrica Metodologia “*Sujeitos*” a descrição dos sujeitos e a métrica Metodologia “*Período*” analisa o período de realização, a métrica “*PalavrasPorSentença*” compõe a categoria Tamanho e Extensão juntamente com as métricas “*Tam Introdução*”, “*Tam Referencial*” e “*Tam Metodologia*”.

As linhas da tabela 7 relativas à “*Aceitou na íntegra*” e “*Recusou na íntegra*” referem-se às vezes em que o professor participante da ferramenta Athos aceitou ou recusou completamente o feedback gerado, de maneira integral, sem alterações, excetuando-se a forma da escrita do feedback.

Ao examinar a tabela 7, nota-se uma frequência reduzida de algumas métricas expressas em valores absolutos. Esse fenômeno decorre principalmente da predominância de feedbacks corretivos em comparação aos elogios aos alunos. Portanto, certas métricas são ativadas apenas quando o sistema identifica erros ou desempenhos abaixo da média. Não há, por outro lado, feedbacks positivos que, por exemplo, informem que o objetivo foi alcançado com uma única ação.

Tanto o gráfico 20 quanto a tabela 7 mostram que os professores participantes aceitaram todas as ocorrências das métricas Objetivo “*AspecMetodoI*”, “*Tam Referencial*” e “*Palavras por Sentença*”, correspondendo a 5,12% das ocorrências. Por outro lado, as métricas “*Título-Objetivo*”, “*Abordagem*” e “*Tam Introdução*”, que correspondem a 1,86% das ocorrências, foram totalmente rejeitadas. Ainda apresenta-se que 38,6% dos trabalhos avaliados tiveram seu feedback aceito na íntegra pelo respectivo professor avaliador e apenas 3,51% dos trabalhos tiveram o feedback totalmente rejeitado.

Esses valores de 38% e 3% para aceitação e rejeição total dos feedbacks, respectivamente, indica uma tendência assertiva por parte da ferramenta Athos na geração das avaliações. Um total de 69% dos feedbacks gerados foram integralmente aceitos pelos professores. Esse valor pode ser considerado alto, uma

vez que qualquer alteração no feedback gerado retira a estatística dessa faixa, colocando-a na faixa de 'alteração' ou 'rejeição'. Na faixa de alteração, muitas vezes o feedback gerado ainda está alinhado à percepção do professor, porém, requer algum ajuste textual. Esses desdobramentos são explorados mais adiante na tese, durante as entrevistas realizadas com os docentes.

Cavalcanti (2021) discute o fato de que o objetivo no uso de ferramentas de correção automatizada não deve ser remover o professor do processo de avaliação, mas sim auxiliá-lo na tarefa. A maneira como a ferramenta Athos foi concebida, sugerindo ao professor feedback avaliativo que pode ser ajustado, está alinhada à visão do pesquisador.

Os resultados obtidos estão também em sintonia com os apresentados por Bastiani (2022). Na pesquisa do autor, foram apresentados valores de concordância de 71%, corroborando a taxa de aceitação deste trabalho que ficou em 69%. A taxa de rejeição dos feedbacks de ambos os trabalhos foi de 3%.

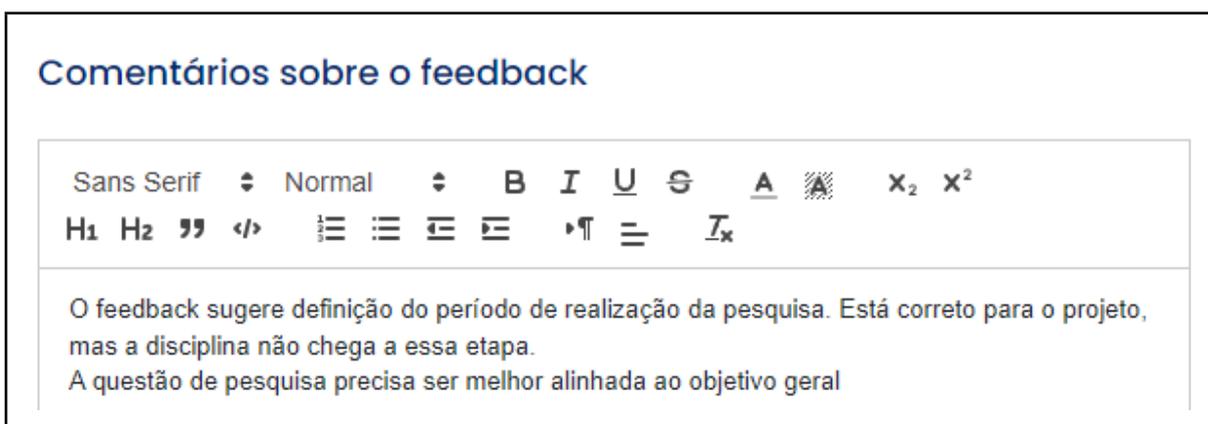
Essa análise dos valores de aceitação ou rejeição está alinhada com os dados coletados através dos comentários e entrevistas realizadas com os professores participantes. Embora a ferramenta Athos gere ocasionalmente feedbacks que não são aceitos, os usuários enfatizam que essas ocorrências são esporádicas e podem ser corrigidas com ajustes direcionados. Todos os participantes destacaram que as ocorrências assertivas são muito mais frequentes, o que é evidenciado pela taxa de aceitação dos feedbacks (69%).

Compreende-se ainda que sempre haverá feedbacks não aceitos ou alterados, uma vez que ferramentas automatizadas não conseguem captar as nuances da escrita de cada aluno, ou mesmo o estilo de comentário de cada professor. Assim, unem-se os benefícios mostrados na literatura ao complementar um feedback automatizado com as considerações e ajustes efetuados pelo professor, integrando os benefícios dos dois estilos como demonstrado por Woodworth (2022). Essa afirmação é corroborada pela aceitação dos feedbacks e fala dos professores participantes.

Ao analisar os comentários dos professores deixados na própria ferramenta, relembra-se que qualquer alteração desses feedbacks gerados fez com que fossem

considerados incorretos para esta análise, mesmo nos casos em que o professor participante apenas removeu o feedback por julgar que a análise realizada não era adequada para a atividade proposta ao aluno. A Figura 17 apresenta um comentário de um professor sobre um feedback gerado que não foi julgado propriamente incorreto, mas inadequado, porque a atividade proposta não considerava aquele aspecto avaliado...

Figura 17 - Comentário sobre o feedback gerado



Fonte: do autor

Na figura 17, o professor participante ressalta que o feedback está correto pois a informação relacionada ao período de realização da pesquisa não está apresentado no trabalho do aluno, porém, o professor ainda informa que a disciplina não chega nesta etapa, ou seja, tal informação não é cobrada do aluno.

Assumindo que um sistema automatizado é composto por conjuntos de regras pré-definidas, torna-se compreensível que esse, ao avaliar um trabalho que não contempla todas as etapas, produza avisos que podem ser considerados incorretos pelo professor. Bird et al. (2009) comentam sobre a necessidade de melhoria constante de sistemas e algoritmos de PLN e isso é coerente neste caso, onde poderia ser dada ao professor a capacidade de escolher quais métricas seriam analisadas, evitando assim esse tipo de consideração de erro.

Outro exemplo disso está na análise dos comentários para a métrica Objetivo “QuantAções”. Nessa métrica obteve-se uma taxa de aceitação de 50%. Ela avalia a existência de mais de um verbo no objetivo da pesquisa. Por vezes o comentário do professor foi favorável, em outros, não, citando que era aceitável a

existência de dois verbos, uma vez que uma ação era complementar à outra, como pode ser visto nas figuras 18 e 19.

Figura 18 - Comentário do professor participante 2 - Métrica “QuantAções”

Objetivo geral

Investigar a cultura afro-gaúcha de Mestre Borel, através da experiência social nos territórios, conhecer os saberes, para visibilidade e importância na construção de experiências contra coloniais na formação de pedagogos.

Comentários sobre o feedback

Sans Serif ▾ Normal ▾ **B** *I* U ~~S~~ **A**  x₂ x²

H₁ H₂ ”” </>       

Destaque para a observação feita no objetivo - ótimo!

Fonte: do autor

A título de recordação, a métrica “QuantAções” produz feedbacks relacionados à quantidade de ações existentes no objetivo da pesquisa do aluno. Para esse caso, o trecho do feedback gerado foi “...mais de uma ação foi identificada no objetivo geral do seu projeto, o que possivelmente indica que ele está dividido em duas ou mais linhas de investigação. Ações identificadas: investigar, conhecer.”

Percebe-se na figura 18 que os verbos utilizados no objetivo estão apresentando mais de uma ação, ou seja, tornam o objetivo confuso e disperso, o que foi reconhecido pelo professor participante 2 no comentário “Destaque para a observação feita no objetivo - ótimo!”.

Figura 19 - Comentário professor participante 1 - Métrica “QuantAções”

Objetivo geral

investigar e avaliar a eficácia do emprego de TIC como ferramentas pedagógicas alternativas no ensino de Biologia Celular, visando aprimorar a compreensão conceitual dos estudantes e promover uma aprendizagem mais significativa e acessível

Comentários sobre o feedback

Sans Serif Normal B I U A x₂ x²

H₁ H₂ ” </> ≡ ≡ ≡ ≡ ¶ ≡

Objetivo com duas ações, mas uma é subsequente a outra, é frequente isso acontecer e maior parte dos examinadores não reclama deste tipo de escrita.

Fonte: do autor

Na figura 19 é possível verificar que os verbos disparadores da métrica Objetivo “QuantAções” são “investigar” e “avaliar” e segundo o comentário do professor participante 1, esse acontecimento pode ser frequente e não é propriamente um erro, uma vez que as ações são complementares uma à outra.

Desta forma, a mesma métrica disparadora recebeu duas avaliações distintas, mesmo não estando propriamente incorreta, apenas sendo sujeita às particularidades de cada situação e avaliador. Esse fator reforça os apontamentos feitos por Hyland e Hyland (2006) ao afirmar que as ferramentas automatizadas devem ser utilizadas em complemento à avaliação efetuada pelo professor, uma vez que professores distintos podem avaliar de maneira diferente certo acontecimento qualitativo, como é o caso dessa métrica. Isso reafirma a necessidade de manutenção do professor no processo avaliativo de projetos de pesquisa.

Assim, contando com 69% de taxa de aceitação das métricas disparadoras e 38% de aceitação integral, nota-se a capacidade dos feedbacks gerados em atenderem às necessidades do professor da disciplina de metodologia da pesquisa, auxiliando-o na correção e avaliação dos trabalhos dos alunos.

Como finalização da leitura do gráfico e da tabela, cabe elencar que as métricas das categorias Sofisticação Léxica, Complexidade Sintática e Dispositivos

de Coesão, assim como as métricas Objetivo “*TermosVagos*” e Metodologia “*PalavrasChave*”, de Estrutura Textual não tiveram ocorrências para os 57 trabalhos avaliados pelos professores participantes. Esse não aparecimento não é visto como problema, uma vez que os trabalhos são desenvolvidos por alunos de nível de graduação e que foram orientados ao longo do semestre por professores de disciplinas de metodologia da pesquisa.

Os resultados estatísticos da aprovação e alteração das validações das extrações automatizadas efetuadas pela ferramenta Athos, avaliados pelos professores participantes estão condensados na tabela 8.

Nessa tabela é possível conferir a quantidade de extrações aceitas pelos professores em cada categoria, título, objetivo e palavras-chave, além da quantidade de alterações efetuadas e o percentual de aceitação para cada categoria. A categoria título do projeto contou com 98,2% de aceitação, sendo que a taxa de inferência da ferramenta de IA foi de 11%. A taxa de inferência é o valor de “títulos” ou “objetivos” extraídos que não existiam de maneira igual no trabalho do aluno. Esses valores demonstram que a capacidade de extração da ferramenta para a categoria Título é alta, inclusive nos casos onde inferências foram apresentadas, tendo o professor validado quase a totalidade das extrações.

Tabela 8 - Resultados da extração automatizada

Validação	Aceitação	Alteração	% Aceitação
Título do projeto	56	1	98,2%
Objetivo do projeto	52	5	91,2%
Palavras-chave	36	21	63,1%

Fonte: do autor.

Para a categoria Objetivo do projeto, a aceitação foi de 91,2%, com uma taxa de inferência de 35,7%. Essa taxa de inferência é compreendida como um ponto positivo das ferramentas de IA e PLN utilizadas, uma vez que os professores participantes validaram a grande maioria das extrações. Essa capacidade positiva é reforçada e evidenciada na figura 20 abaixo, onde o professor participante pontua

que o objetivo apresentado pela ferramenta não está idêntico ao descrito pelo aluno, o que o alerta para verificar qual das duas formas é a mais adequada.

A categoria Palavras-chave, por sua vez, foi a que obteve menor taxa de aceitação por parte dos professores participantes da ferramenta, resultando em 63,1% de validação das palavras-chave extraídas. A esse resultado são atribuídas as particularidades de identificação obtidas através das ferramentas de PLN e IA utilizadas, as quais não possuem as capacidades de um professor na identificação de pontos internos e específicos.

Figura 20 - Comentário sobre apresentação do objetivo.

Comentários sobre o feedback

Sans Serif ↕ Normal ↕ B I U ☒
A 🔍 x_2 x^2 H_1 H_2 ” </>
 $\frac{1}{3}$ ≡ ⇐ ⇒ ↗ ≡ $\frac{1}{x}$

Comentário positivo sobre a extração de objetivo geral: a ferramenta está extraíndo o objetivo, muitas vezes eliminando partes do que o aluno escreve que são efetivamente irrelevantes. Está escrevendo a essência do objetivo. O fato de deixar marcado que não é assim que aparece no texto, pode ajudar o professor a localizar se a maneira como o aluno elaborou o objetivo estava clara ou não.

Fonte: do autor

Com esses valores de aceitação, conclui-se que a ferramenta Athos tem capacidade de identificar corretamente o título e o objetivo da pesquisa do aluno, obtendo taxas de acerto superiores a 91%. Para o elemento palavras-chave, apesar da taxa de aceitação de 63%, reforça-se a necessidade de validação desse por parte do professor, uma vez que esse elemento é utilizado em processos da ferramenta. Ao mesmo tempo, reconhece-se esse como um ponto para melhoria da ferramenta, utilizando-se de novas técnicas ou ferramentas para a extração e identificação dessas palavras-chave.

A última fase do estudo apoiou-se em uma abordagem qualitativa. Após o uso do sistema para avaliação dos trabalhos desenvolvidos por seus alunos nas respectivas disciplinas, os professores passaram por uma entrevista para que se pudesse conhecer suas percepções sobre o uso da ferramenta Athos. Esses dados foram submetidos a técnicas de Análise de Conteúdo por dois codificadores diferentes, sendo os códigos determinados como disposto na tabela 9 abaixo.

Tabela 9 - Códigos da Análise de Conteúdo

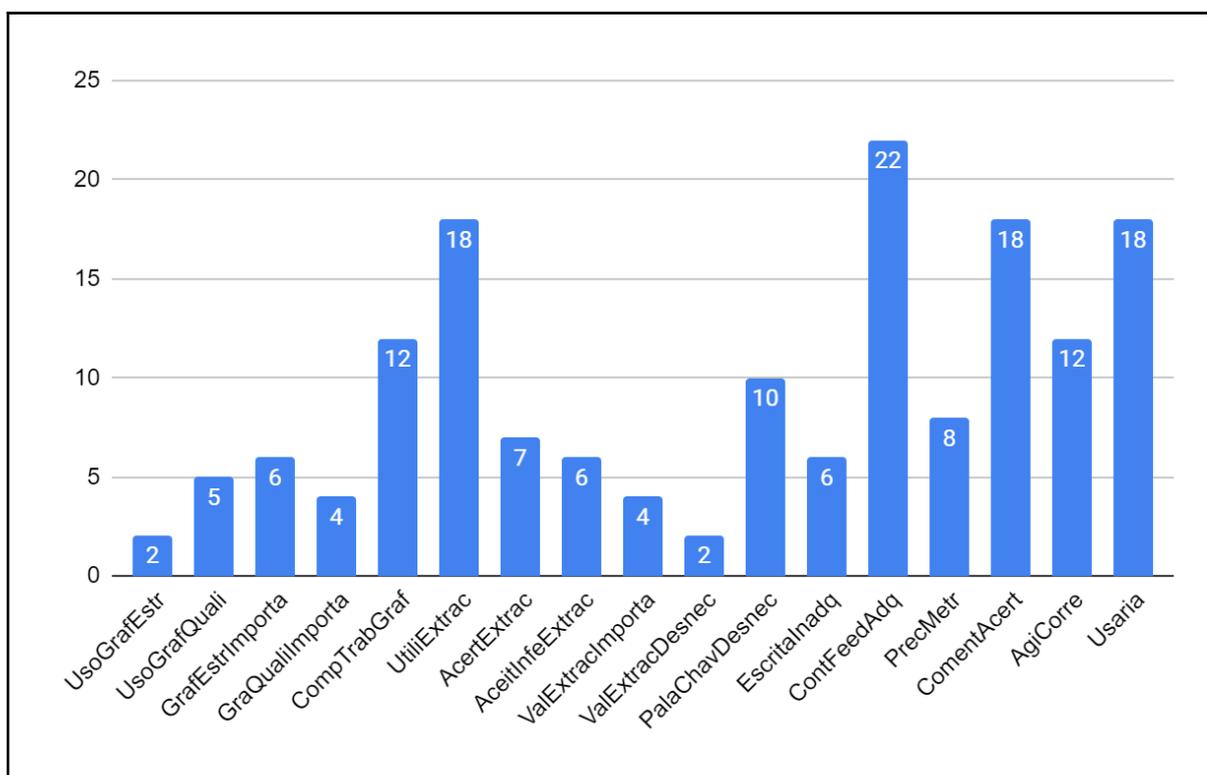
UsoGrafEstr	O professor utilizou o gráfico de Estrutura Textual?
UsoGrafQuali	O professor utilizou o gráfico de Complexidade Textual?
GrafEstrImporta	O professor considera o gráfico de Estrutura importante?
GraQualiImporta	O professor considera o gráfico de Complexidade importante?
CompTrabGraf	O professor gostou da comparação entre os trabalhos apresentados pelos gráficos?
UtiliExtrac	A extração dos componentes foi útil ao professor?
AcertExtrac	A extração estava precisa ou o professor precisou corrigir muitas vezes?
AceitInfeExtrac	O professor aceitou bem as inferências feitas durante a extração?
ValExtracImporta	O professor julga importante a validação da extração?
ValExtracDesnec	O professor julga desnecessária a validação da extração?
PalaChavDesnec	O professor considera ter as palavras chave em vista como irrelevante?
Escritalnadq	A forma da escrita do feedback estava inadequada?
ContFeedAdq	O conteúdo do feedback estava adequado?
PrecMetr	Existem métricas que podem receber ajustes nos parâmetros?
ComentAcert	O professor avalia os feedbacks fornecidos pela ferramenta como assertivos?
AgiCorre	O sistema agilizou o trabalho do professor?
Usaria	O professor utilizaria a ferramenta consistentemente em todas as ofertas da disciplina?

Fonte: do autor

A fala e os comentários dos professores foram analisados através do processo de codificação da análise de conteúdo proposto por Saldaña (2013). Para essa tarefa foram convidados dois pesquisadores com experiência na área, para que, de modo independente, identificassem códigos dos temas tratados nos dados coletados. Do resultado desta codificação foi extraído o nível de concordância entre os codificadores, conferido pelo cálculo do coeficiente Cohen's Kappa, que classifica

os níveis de concordância entre “Pequeno” (0.00) e “Quase Perfeito” (1.00). O cálculo resultou em um coeficiente de valor 0.81, obtendo classificação igual ou superior à “Significativa”. O percentual de concordância entre os codificadores foi de 85,7%.

Figura 21 - Quantidade de ocorrência dos códigos



Fonte: do autor

A figura 21 acima mostra o número de ocorrência de cada código. Em uma primeira análise, destacam-se os códigos com menor número de ocorrências, ao exemplo de UsoGrafEstr, GraQualiImporta, ValExtracImporta e ValExtracDesnec. Esses códigos tiveram apenas uma ocorrência por professor, não havendo consenso entre eles. No caso do código UsoGrafEstr, o professor participante 1 citou que fez uso do gráfico e que ele é importante para uma rápida identificação de qual trabalho possui mais erros e possivelmente demandará mais trabalho e atenção. Já os professores participantes 2 e 3 citaram que não utilizaram o gráfico de estrutura por demandas de tempo e preferência em acessar o trabalho diretamente.

No caso dos códigos ValExtracImporta e ValExtracDesnec, os professores participantes divergiram na questão da importância da validação dos dados extraídos de forma automatizada pela ferramenta Athos, a saber: título do trabalho,

objetivo geral e palavras-chave. O professor participante 1 citou a compreensão da importância da validação para a ferramenta, mas que esta ação de validar os dados extraídos automaticamente, para ele, na ação de avaliador do trabalho do aluno, é desnecessária. Já o professor participante 2 cita que esta ação de validar os dados extraídos exige e reforça os conceitos do trabalho do aluno ao professor, fazendo com que este esteja mais atento a estes pontos durante a correção. O professor participante 3 comenta que considera importante a ação de validação, pois dá ao professor a oportunidade de corrigir possíveis erros da extração automatizada.

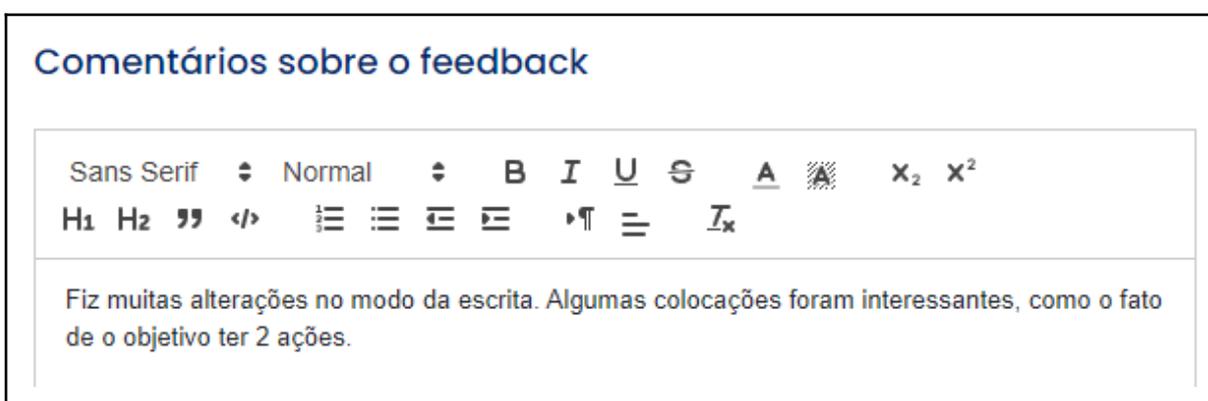
Esta divergência de opiniões quanto à necessidade de validação da extração automatizada é associada às visões do professor participante com o funcionamento da ferramenta desenvolvida. Enquanto o professor 1 reconhece a necessidade da ação para o funcionamento da ferramenta, os professores 2 e 3 não fazem esta vinculação. Avançando na visão da necessidade de validação da extração automatizada, os trabalhos de Brown et al. (2020), Radford et al. (2019) e também a própria mantenedora do Chat GPT informam que as ferramentas de IA generativa podem cometer erros ou apresentar dados incorretos, sendo que pontos importantes devem ser checados e conferidos.

Continuando nos códigos sobre a extração automatizada da ferramenta, o código UtiliExtrac refere-se à utilidade e importância que o professor participante dá ao fato de ter estes dados extraídos pela ferramenta. Os utilizadores citaram e reforçaram durante suas falas a importância desta extração.

A figura 22 abaixo mostra um comentário do professor participante 1 sobre o tema.

Quanto ao código Escritalnadq, todos os utilizadores da plataforma citaram a necessidade de reestruturar a forma como os feedbacks foram enviados ao aluno. Isto se deve ao fato de que cada métrica disparadora fornece um feedback sobre seu tema, no final da geração, todas as frases de feedback são enviadas à ferramenta de IA generativa, Chat GPT, para que esta transcreva as frases em um texto corrido e coeso. É sobre esta transcrição que os utilizadores apresentaram queixas, citando por exemplo que ou a escrita era muito formal, como “Prezado aluno fulano de tal, parabenido vossa senhoria...”, o que não condizia com a forma com a qual o respectivo professor participante se comunica com seus alunos.

Figura 23 - Comentário sobre alterações no modo da escrita



Fonte: do autor

Além da figura 23, que apresenta o comentário do professor participante indicando que as colocações da ferramenta foram interessantes, durante a entrevista o professor participante 2 também citou que “*os termos da ferramenta não foram um problema, mas preferi utilizar meus termos*”. Outro comentário interessante levanta o ponto de que cada professor possui uma forma específica e diferente para comunicar-se com seus alunos, o que inviabiliza a capacidade da ferramenta em reproduzir esta forma de escrita de maneira adequada. A fala foi: “*não vale a pena o sistema copiar o meu padrão de escrita, o meu padrão pode ser diferente de outro professor e quanto isto custaria para o sistema?*”

Considerando o uso de IA generativa, mais especificamente o Chat GPT, para a conversão dos feedbacks individuais em um texto corrido e coeso, atribui-se o problema reportado pelos professores participantes como um problema de “prompt”, ou seja, da forma como a solicitação de conversão do texto foi solicitada à

ferramenta de IA. Este assunto da construção do prompt de comando da IA é tratado no trabalho de Arora et. al. (2023), os autores comentam sobre os poderes de entendimento da ferramenta de IA e como pequenas alterações na solicitação geram uma resposta totalmente diferente.

O resultado final do texto de feedback foi apontado pelos professores como um problema, porém algo de pequena dimensão. Desta forma, assume-se este problema como algo a ser melhorado em versões ou trabalhos futuros, onde deverão ser analisados diversos comandos e solicitações, assim como as respectivas respostas da ferramenta de IA generativa.

Outros dois códigos importantes são o AgiCorre, que refere-se ao quanto o professor considera que a ferramenta agiliza a correção dos trabalhos e o código Usaria, que questiona a vontade do professor em utilizar a ferramenta de forma contínua em suas disciplinas de metodologia da pesquisa.

Para o código AgiCorre, os três professores concordaram que a utilização do sistema agiliza e facilita a correção dos trabalho, mesmo que uma análise quantitativa de tempo despendido na atividade possa apresentar valores maiores do que sem o uso da ferramenta, o professor participante 2 comenta que *“...agiliza sim, tomou mais trabalho pois tinha que avaliar o aluno e retornar e dar o comentário na ferramenta...”*. Já o professor participante 1 coloca que *“...claro que sempre uma ferramenta nova, ela tem um tempo de aprendizagem. Né, então no início a gente demora mais...”*. Este comentário é complementado com: *“...mas depois que a pessoa entende como o sistema funciona, certamente ele agiliza, ele ajuda a acelerar mesmo o processo de revisão”*, o professor participante 3 salienta que *“eu utilizei a ferramenta depois de corrigir os trabalhos dos alunos, mas é clara a capacidade do sistema em agilizar o trabalho”*.

O entendimento de que uma ferramenta computacional pode auxiliar e agilizar o trabalho do professor é abordado na literatura, como em Jurafsky e Martin (2018) e em Zhai e Ma (2022). Este entendimento foi corroborado pela fala dos professores participantes através de diversos comentários sobre a ferramenta Athos, sendo unânime os apontamentos das capacidades positivas da ferramenta em agilizar o trabalho de correção, como por exemplo ao apresentar os elementos do trabalho de forma facilitada, elencar inconsistências entre a extração automatizada e

o texto do aluno ou diretamente no fornecimento de feedbacks direcionados a critérios específicos.

O código Usaria foi marcado por comentários como “...gostaria de utilizar sim, principalmente se houver maior integração com o MOODLE...”, “...mesmo com projetos preliminares já dá retorno...”, “...eu vejo um enorme potencial na ferramenta para apoiar a revisão de projetos da maneira como ela tá construída, mas entendê-la então para auxiliar professores na revisão e acompanhamento né de dissertações e teses...” e “...gostaria muito de usar no próximo semestre, pois terei bem mais alunos e a ferramenta auxilia muito o trabalho...”.

Estes apontamentos permitem compreender que a ferramenta Athos auxiliou os professores na tarefa de correção e avaliação de projetos de pesquisa de seus alunos, fornecendo comentários assertivos e agilizando o processo de avaliação dos trabalhos.

4 CONCLUSÕES

Os estudos que compuseram esta tese permitiram concluir que a ferramenta de análise textual desenvolvida, baseada em técnicas que combinaram métricas da linguística computacional com métricas de estrutura textual, foi capaz de auxiliar os docentes na atividade de acompanhamento e avaliação de projetos de pesquisa em disciplinas de graduação.

Os estudos abordaram análises de redações e projetos de pesquisa sob a tutela de métricas de complexidade textual, correlacionadas com as notas atribuídas por avaliadores especialistas.

Os resultados apresentados demonstram as capacidades destas métricas textuais em predizer ou indicar possíveis notas atribuídas aos trabalhos, fornecendo embasamento para novas ferramentas avaliativas que têm como meta auxiliar professores em correções de trabalhos textuais.

Observou-se também, por meio dos resultados dos estudos desenvolvidos, que diferentes estilos textuais afetam e alteram a eficácia das métricas. Isto é demonstrado nas diferenças obtidas entre os estudos de redações e de projetos de pesquisa. No estudo de redações, os resultados permitem concluir que as métricas não são estatisticamente suficientes para a predição da nota atribuída pelo avaliador. Estas estatísticas diferem dos resultados reportados na literatura para outros estilos textuais, como os expressados no estudo de projetos de pesquisa.

Para os projetos de pesquisa, as métricas textuais obtidas se mostraram estatisticamente significativas e correlacionadas com as notas atribuídas pelo professor da disciplina, o que demonstra que o gênero textual tem influência sobre o resultado da correlação métrica/nota. Assim, vislumbra-se a necessidade de maiores análises envolvendo diferentes tipos textuais e as métricas textuais.

Foram mineradas, selecionadas e testadas métricas de complexidade e estrutura textual para formarem um conjunto de métricas de mensuração para a geração de feedbacks avaliativos. Este conjunto foi testado e avaliado nos estudos que compõem esta tese, nos quais foram encontradas métricas estatisticamente significativas para a tarefa, permitindo a construção de níveis e parâmetros de uso e

conversão destas métricas em feedbacks textuais que auxiliassem o professor na tarefa de avaliação e correção de projetos de pesquisa.

A validação dos feedbacks, das métricas e seus parâmetros de geração foi efetuada com o uso da ferramenta por três professores de disciplinas de metodologia da pesquisa de nível de graduação. Esta validação contou com a análise de 57 trabalhos de pesquisa, onde foram obtidos pareceres dos professores com relação aos feedbacks gerados. Os feedbacks foram aceitos 69,4% das vezes em que foram disparados, sendo que apenas 30,6% dos feedbacks receberam alguma alteração por parte do professor. Considerando-se que cada texto final é composto pelo disparo de várias métricas e cada métrica relacionada a um feedback exclusivo, este valor de 69,4% de aceitação é tido como positivo. Este fator de sucesso é ainda maior ao observar-se que 38% dos textos finais aceitos sem modificação e apenas 3,5% totalmente rejeitados.

O desenvolvimento de uma ferramenta tecnológica para avaliação automatizada de projetos de pesquisa se mostrou um desafio, diversas técnicas e algoritmos de PLN e IA generativa, juntamente com ferramentas de mineração, manipulação e identificação textual foram utilizadas para permitir esta construção. Ela contou com diversas etapas de testes e validações até o refinamento do processo e a construção final da ferramenta.

Mesmo assim, durante a fase de entrevistas e validação da ferramenta junto dos professores participantes, foram elencadas limitações e melhorias que podem tornar a ferramenta Athos ainda mais competente em seu objetivo, seja por analisar diferentes aspectos ou desenvolver novas capacidades.

Dentre estas limitações e demandas, salientam-se algumas como:

- Detector de plágio;
- Correção de formatação;
- Conferência de citações e uso de referências;
- Ajustes na geração do texto final de feedback;
- Analisar a etapa de resultados da pesquisa.

Entende-se que estas demandas podem enriquecer e aumentar o potencial avaliativo da ferramenta, adicionando novas capacidades de análise e aumentando

o alcance das já existentes. Sendo bons pontos para trabalhos e atualizações futuras.

4.1 Considerações Finais

Este trabalho permitiu aprofundar os conhecimentos acerca da escrita e avaliação de trabalhos de pesquisa, explorando áreas da Analítica da Aprendizagem juntamente com técnicas de PLN. Foram apresentadas ferramentas capazes de auxiliar no processo de obtenção de insights na identificação de padrões textuais, além de fornecer uma ferramenta inovadora para o auxílio do professor na correção e avaliação de projetos de pesquisa, juntamente com seus corpus base.

Os achados deste trabalho estão em sintonia com trabalhos relatados na literatura. Assim como nos trabalhos de Crossley e McNamara (2011) e McNamara e Graesser (2012), ficou claro que o uso de métricas textuais pode auxiliar em diversas identificações e complementar a avaliação do professor, tornando esta avaliação mais rápida e precisa.

A ferramenta desenvolvida mostrou-se válida na atividade proposta, tendo aceitação dos professores participantes e mostrando resultados promissores no auxílio destes, permitindo que concentrem sua avaliação em pontos e particularidades de cada aluno.

Os resultados obtidos demonstram a capacidade da ferramenta em avaliar projetos de pesquisa e fornecer feedbacks auxiliares aos professores, obtendo taxas de 69% de aceitação nos feedbacks gerados. O que está em sintonia com pesquisas apresentadas.

Além disso, este trabalho foi capaz de responder a questão norteadora da pesquisa, apresentando como uma ferramenta computacional com funções de avaliação automatizada da escrita pode auxiliar o professor na avaliação e elaboração de feedbacks sobre textos do gênero projeto de pesquisa.

Para esta apresentação, utilizou-se de técnicas de PLN, sistemas de extração, avaliação e mineração textual e integração com ferramentas de IA e ambientes virtuais de aprendizagem. Todas estas ferramentas culminaram na

construção e desenvolvimento da ferramenta Athos, que através de um portal web gratuito, embasada em métricas de análise da estrutura textual e da complexidade textual, fornece feedbacks avaliativos e formativos que auxiliam o professor na tarefa de correção e avaliação de projetos de pesquisa.

A ferramenta Athos conta com uma grande gama de trabalhos analisados previamente, que compõem seu corpus de conhecimento, além de estar embasada em trabalhos atuais da área para a construção das regras sobre os elementos que compõem um projeto de pesquisa, bem como a maneira que tais elementos se relacionam dentro do texto.

Como forma de validação da ferramenta, esta foi submetida a uso em ambiente real, por professores de disciplinas de metodologia científica, o que tornou a coleta de dados fidedigna e completa, tendo os resultados evidenciado a importância da aplicação das tecnologias no contexto educacional.

Pelo lado tecnológico, a ferramenta Athos mostrou-se capaz de atuar na avaliação e geração de feedbacks sobre projetos de pesquisa, identificando fragilidades e pontos de melhoria nos trabalhos, apresentando taxas adequadas de aceitação e concordância com os professores participantes.

Esta pesquisa abre portas, tanto para continuidade deste quanto para novos estudos na área da Analítica da Escrita. As diferenças encontradas entre os corpora da pesquisa corroboram a necessidade de estudos mais específicos e aprofundados para cada área de ensino. Além disso, as diversas métricas textuais obtidas com ferramentas de PLN demandam refinamento constante e extensa aplicação nas diversas formas textuais.

Um trabalho futuro pode ser a inclusão de sistemas de recomendação de conteúdo e a possibilidade de ampliação de uso da ferramenta, contemplando na totalidade os gêneros de elementos comuns, como artigos, dissertações e teses, além da aplicabilidade das métricas e feedbacks para outras áreas de ensino.

Por fim, vale ressaltar que o desenvolvimento de soluções tecnológicas para apoio da educação, em especial à produção textual, são de grande importância para a sociedade e o desenvolvimento humano. Em um ambiente cada vez mais rápido e ágil, é fundamental que o professor tenha ferramentas que o auxiliem nos processos

cotidianos da profissão, permitindo que este concentre-se no aluno e suas particularidades.

Neste sentido, esta pesquisa utilizou de tecnologias, modelos e estratégias pedagógicas para auxiliar professores no processo avaliativo da escrita acadêmica, o que tem bom potencial para trazer futuramente repercussões educacionais e sociais positivas.

REFERÊNCIAS

ADAM, Bruna C. N., BENITES, Maure C. dos S., FLORIANO, Mikaela D. P., ADÃO, Sebastião A. da R. OS DESAFIOS DA CARREIRA DOCENTE NO ENSINO SUPERIOR: UM ESTUDO COM PROFESSORES DA UNIVERSIDADE FEDERAL DO PAMPA. **XIX Colóquio Internacional de Gestão Universitária**. 2019. Florianópolis. Disponível em: https://repositorio.ufsc.br/bitstream/handle/123456789/201902/101_00170.pdf?sequence=1&isAllowed=y. Acesso em: 20 de mar de 2023.

ALUÍSIO, Sandra M., PELIZZONI, Jorge M., MARCHI, Ana. R., de OLIVEIRA, Lucélia. H., Manenti, Regiana, MARQUIAFÁVEL, Vanessa, TELES, Jorge. An account of the challenge of tagging a reference corpus for brazilian portuguese. **Computational Processing of the Portuguese Language**. Tradução . Heidelberg: Springer, 2003. Disponível em: https://repositorio.usp.br/directbitstream/fe81be94-efbc-4ac7-969f-193faf4d0d2e/Relat%C3%B3rio%20T%C3%A9cnico_188_2.... Acesso em 05 de dez de 2023.

ALUISIO, Sandra et al. Readability Assessment for Text Simplification. In: NAACL HLT, 2010, Los Angeles, CA. **Fifth Workshop on Innovative Use of NLP for Building Educational Applications**. Los Angeles, CA: Association for Computational Linguistics, 2010. p. 1–9. Disponível em: <https://aclanthology.org/W10-1001.pdf>. Acesso em: 29 out. 2023.

ARAÚJO, Júlio; DIEB, Messias; COSTA, Sayonara Melo. O QNP e as dificuldades de construção do objeto de pesquisa: uma experiência de aprendizagem mediada sobre o gênero projeto de pesquisa. **DELTA: Documentação de Estudos em Linguística Teórica e Aplicada**, [S. l.], v. 33, n. 3, p. 729–757, 2017. DOI: 10.1590/0102-445097873044367046.

ARORA, Simran et. al. Ask Me Anything: A simple strategy for prompting language models. In: EICLR, 2023. **The Eleventh International Conference on Learning Representations**. Disponível em <https://openreview.net/forum?id=bhUPJnS2g0X>. Acesso em 03 mai 2024.

ASSIS, Juliana Alves. Ações do professor e do universitário nas práticas de ensino e de aprendizagem da escrita acadêmica: o papel da avaliação e da reescrita no processo de apropriação do gênero resenha. **Eutomia - Revista de Literatura e Linguística**. v. 1, n. 13, p. 543–561, 2014. b. DOI: 10.19134/eutomia-v1i13p543-561.

BARBOSA, Vanessa Fonseca. **Uma voz apagada? : análise da atividade de revisão de textos acadêmicos sob as perspectivas bakhtiniana e ergológica**. 2017. 204f. Tese (Doutorado em Linguística) - Pontifícia Universidade Católica do Rio Grande do Sul – PUCRS. Programa de Pós-Graduação em Letras. Acesso em 05 de abr de 2023. Disponível em: <https://tede2.pucrs.br/tede2/handle/tede/7865> .

BASTIANI, Ederson. **Desenvolvimento e validação de uma ferramenta computacional apoiada por mineração de texto voltada à qualificação da escrita de projetos de pesquisa**. Tese (Doutorado em Informática na Educação) Universidade Federal do Rio Grande do Sul - Programa de Pós-Graduação em Informática na Educação - 2022. 119 f. Disponível em: <https://lume.ufrgs.br/bitstream/handle/10183/249118/001150010.pdf?sequence=1&isAllowed=y>
Acesso em 14 de ago de 2023

BAZERMAN, C. What written knowledge does: Three examples of academic discourse. **Philosophy of the Social Sciences**, 11(3), 361-387. 1981

BAZERMAN, C. The life of genre, the life in the classroom. In Bishop, W. e Ostrom, H. (Eds.), **Genre and Writing: Issues, Arguments, Alternatives** (pp. 19-26). Boynton/Cook Publishers. 1997

BAZERMAN, C. (2004). Speech Acts, Genres, and Activity Systems: How Texts Organize Activity and People. In Freedman A. e Medway P. (Eds.), **Genre and the New Rhetoric** (pp. 23-36). Routledge. 2004

BENDER, Emily M. e FRIEDMAN, Batya. Data Statements for Natural Language Processing: Toward Mitigating System Bias and Enabling Better Science. **Transactions of the Association for Computational Linguistics**, 6:587–604. 2018. Acesso em 15 de nov de 2023. Disponível em: <https://aclanthology.org/Q18-1041.pdf> >

BENGIO, Y., COURVILLE, A., E VINCENT, P. Representation Learning: A Review and New Perspectives. **IEEE Transactions on Pattern Analysis and Machine Intelligence**. 2013

BEREITER, Carl. e SCARDAMALIA, Marlene. **The psychology of written composition**. Routledge. 1987. Disponível em: <https://www.routledge.com/The-Psychology-of-Written-Composition/Bereiter-Scardamalia/p/book/9780805800388>> Acesso em: 15 de mar. de 2023

BEREITER Carl e SCARDAMALIA Marlene. An Attainable Version of High Literacy: Approaches to Teaching Higher-Order Skills in Reading and Writing, **Curriculum Inquiry**, 17:1, 9-30. 1987. DOI: 10.1080/03626784.1987.11075275. Disponível em: <https://www.jstor.org/stable/1179375>> Acesso em: 15 de mar. de 2023

BICK, E. **The Parsing System PALAVRAS**. Aarhus University, Denmark. 2000. Acesso em 05 de dez de 2023. Disponível em: <https://edu.visl.dk/~eckhard/pdf/PLP20-amilo.ps.pdf>>

BIRD, S., LOPER, E. e KLEIN, E. **Natural Language Processing with Python**. O'Reilly Media, Inc. 2009

BOOTH, Wayne C.; COLOMB, Gregory G.; WILLIAMS, Joseph M. **The Craft of Research**. 4. ed. Chicago: The University of Chicago Press, 2016. Acesso em 20 de dez de 2023. Disponível em: <<https://press.uchicago.edu/ucp/books/book/chicago/C/bo23521678.html>>

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). A Redação do Enem 2023: cartilha do participante. Brasília, 2023. Disponível em: https://download.inep.gov.br/publicacoes/institucionais/avaliacoes_e_exames_da_educacao_basica/a_redacao_no_enem_2023_cartilha_do_participante.pdf

BROWN, T. B. et al. Language Models are Few-Shot Learners. 2020. OpenAI. <https://doi.org/10.48550/arXiv.2005.14165>

BRYMAN, A. Social research methods. Oxford University Press 2016.

BRUTON Samuel e CHILDERS Dan. The ethics and politics of policing plagiarism: a qualitative study of faculty views on student plagiarism and Turnitin®. **Assessment & Evaluation in Higher Education**, 41:2, 316-330. 2016. DOI: 10.1080/02602938.2015.1008981 Acesso em 27 de mar de 2023. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/02602938.2015.1008981?scroll=top&needAccess=true&role=tab> .

BURSTEIN, J., CHODOROW, M. e LEACOCK, C. Automated Essay Evaluation: The Criterion Online Writing Service. **AI Magazine**, 25(3), 27. 2004.. <https://doi.org/10.1609/aimag.v25i3.1774>

CAMPOS, A. de; GALAFASSI, C.; BASTIANI, E.; JOSENDE PAZ, F.; LEITE CAMPOS, R.; KRUG WIVES, L.; CESAR CAZELLA, S.; BERNI REATEGUI, E.; AUGUSTO COUTO BARONE, D. Mineração de Dados Educacionais e Learning Analytics no contexto educacional brasileiro: um mapeamento sistemático. **Informática na educação: teoria & prática**, Porto Alegre, v. 23, n. 3 Set/Dez, 2020. DOI: 10.22456/1982-1654.102618. Disponível em: <<https://seer.ufrgs.br/index.php/InfEducTeoriaPratica/article/view/102618>>. Acesso em: 28 jun. 2023.

CARLESS, D. Learning-oriented assessment: Principles and practice. **Assessment & Evaluation in Higher Education**, 32(3), 267-286. 2007 .DOI:[10.1080/02602930600679043](https://doi.org/10.1080/02602930600679043)

CARLESS, D. Excellence in university assessment: Learning from award-winning practice. Routledge. 2015.

CARLESS, D.; BOUD, D. The development of student feedback literacy: enabling uptake of feedback. **Assessment & Evaluation in Higher Education**, v. 43, n. 8, p. 1315-1325, 2018. Acesso em: <20 de

fevereiro de 2023> Disponível em:
 <<https://www.tandfonline.com/action/showCopyRight?scroll=top&role=tab&doi=10.1080%2F02602938.2018.1463354>>

CASNICI, N. e GRIMALDO, F. How to peer review a paper: a guide for newcomers and a refresher for the experienced. **Research Integrity and Peer Review**, 4(1), 14. 2019.

CAVALCANTI, A. P. et al. Automatic feedback in online learning environments: A systematic literature review. **Computers and Education: Artificial Intelligence**, v. 2, p. 100024, 2021. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2666920X21000217> . Acesso em: 05 fev. 2023.

CHING, Yu-Hui e HSU, Yu-Chang. "Peer Feedback to Facilitate Project-Based Learning in an Online Environment." **International Review of Research in Open and Distributed Learning**, volume 14, number 5, december 2013, p. 258–276. <https://doi.org/10.19173/irrodl.v14i5.1524> Acesso em 05 de abr de 2023. Disponível em: <https://www.erudit.org/en/journals/irrodl/1900-v1-n1-irrodl05093/1066893ar/abstract/> .

CHURCHILL, G. A., IACOBUCCI, D. e GLEN, W. J.. Marketing research: Methodological foundations. **Cengage Learning**. 2018.

CRESWELL, J. W. Research design: Qualitative, quantitative, and mixed methods approaches. Sage Publications. 2014.

CROSSLEY, S. A., MCNAMARA, D. S. e MCCARTHY, P. M. Learning analytics and writing instruction. **Computers and Composition**, 27(1), 34-47. 2010.

CROSSLEY, S. A., ALLEN, D. B. e MCNAMARA, D. S. Text readability and intuitive simplification: A comparison of readability formulas and Coh-Metrix. **Reading in a Foreign Language**, 23(1), 84-101. 2011. Acesso em 6 abr 2023. Disponível em: <https://files.eric.ed.gov/fulltext/EJ926371.pdf>

COPE, B. e KALANTZIS, M. Big Data Comes to School: Implications for Learning, Assessment, and Research. **AERA Open**, 2(2). 2016. <https://doi.org/10.1177/2332858416641907>

CONNEAU, A. et al. Unsupervised Cross-lingual Representation Learning at Scale. Facebook AI Research. 2020. <https://doi.org/10.48550/arXiv.2006.13979>

DAWSON, S., GASEVIC, D., SIEMENS, G. e JOKSIMOVIC, S. Current state and future trends: A citation network analysis of the learning analytics field. In Proceedings of the 8th International Conference on Learning Analytics and Knowledge (pp. 231-240). 2018. ACM.

DEVLIN, J., CHANG, M. W., LEE, K. e TOUTANOVA, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proceedings of NAACL-HLT. 2018. Acesso em 02 de dez de 2023. Disponível em:

<https://aclanthology.org/N19-1423.pdf?utm_medium=email&utm_source=transaction>

EPSTEIN, D.; REATEGUI, E. Uso de mineração de textos no apoio à compreensão textual. RENOTE, Porto Alegre, v. 13, n. 1, 2015. DOI: 10.22456/1679-1916.57647. Disponível em: <https://www.seer.ufrgs.br/index.php/renote/article/view/57647>. Acesso em: 10 maio. 2023.

EVERS, Aline; FINATTO, Maria José Bocorny. Linguística de Corpus, Léxico-Estatística Textual e Processamento de Linguagem Natural: perspectiva para estudos de vocabulário em produções textuais. Revista GTLex, [s. l.], v. 1, n. 2, p. 271, 2016. Disponível em: <https://doi.org/10.14393/lex2-v1n2a2016-3>. Acesso em: 29 out. 2023.

FLOWER, L. e HAYES, J. R. A Cognitive Process Theory of Writing. College Composition and Communication, [S. l.], v. 32, n. 4, p. 365, 1981. DOI: 10.2307/356600.

GALLIEN, Tara e EARLY, Jody. Personalized Versus Collective Instructor Feedback in the Online Courseroom: Does Type of Feedback Affect Student Satisfaction, Academic Performance and Perceived Connectedness With the Instructor?. International Journal on E-Learning. 7. 2005.

GARCIA, L. P.; BENEVIDES-PEREIRA, A. M. T. Investigando o Burnout em Professores Universitários. [S. l.], v. 18, n. 1, p. 76–89, 2005.

GIL, Antonio Carlos. Como elaborar projetos de pesquisa. [s.l: s.n.]. DOI: 10.1007/978-1-137-49662-1.

GRAESSER, A. C., MCNAMARA, D. S., LOUWERSE, M. M. e CAI, Z. Coh-Metrix: Analysis of text on cohesion and language. **Behavior Research Methods, Instruments, & Computers**, 36(2), 193-202. 2004. Acesso em: 06 abr 2023. Disponível em: <https://link.springer.com/article/10.3758/bf03195564> .

HATTIE, John; TIMPERLEY, Helen. The Power of Feedback. Review of Educational Research, v. 77, n. 1, p. 81-112, 2007. DOI: <https://doi.org/10.3102/003465430298487>. Disponível em: <http://www.columbia.edu/~mvp19/ETF/Feedback.pdf> . Acesso em: 17 mar. 2023.

HATTIE, John. Visible Learning. London, England: Routledge. 2008.

HAYES, J.R., ROWER, L.S., SCHRIVER, K.A., STRATMAN, J. e CAREY, L.. Cognitive processes in revision. In S. Rosenberg (Ed.). Advances in psycholinguistics. Vol. 2: Reading, writing and language (pp. 176-240). Cambridge: Cambridge University Press. 1987

HAYES, John R.; FLOWER, Linda S. Identifying the Organization of Writing Processes. *Cognitive Processes in Writing*, [S. l.], n. December, p. 5–30, 1980. DOI: BF 456 W8 C676 1980.

HIGGINS, R., HARTLEY, P. e SKELTON, A. The conscientious consumer: Reconsidering the role of assessment feedback in student learning. *Studies in Higher Education*, 27(1), 53-64. 2002. <https://doi.org/10.1080/03075070120099368>

JUGRAN, S., KUMAR, A., TYAGI, B. S. e ANAND, V. Extractive automatic text summarization using SpaCy in Python & NLP. In 2021 International conference on advance computing and innovative technologies in engineering (ICACITE) (pp. 582-585). 2021. DOI: 10.1109/ICACITE51222.2021.9404712

HOVY, E. e LAVID, J. Toward a 'Science' of Corpus Annotation: A New Methodological Challenge for Corpus Linguistics. **International Journal of Translation**. v. 22, n. 1, p. 13-36, 2010.

HYLAND, K., HYLAND, F. Feedback on Second Language Students' Writing. *Language Teaching* 39:83–101, 2006. doi:10.1017/S0261444806003399. Acesso em 15 de jun de 2023. Disponível em: <<https://www.cambridge.org/core/journals/language-teaching/article/abs/feedback-on-second-language-students-writing/2A58FAA099F7628B459AE677F7098493>>

HUTCHINS, W. J. *The History of Machine Translation*. Routledge. 2006. DOI: <https://doi.org/10.1016/B978-0-08-042580-1.50066-0>

JURAFSKY, D. e MARTIN, J. H. "Speech and Language Processing" (3rd ed.). 2018. Acesso em 15 de nov. de 2023. Disponível em: <<https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf>>

KLEIN, D. e MANNING, C. D. Accurate Unlexicalized Parsing. *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics*. p. 423-430. 2003.

KNIGHT, S., SHIBANI, A., ABEL, S., GIBSON, A., RYAN, P., SUTTON, N., WIGHT, R., LUCAS, C. AcaWriter: A learning analytics tool for formative feedback on academic writing. *Journal of Writing Research*, 12(1), 141–186. 2020. <https://doi.org/10.17239/jowr-2020.12.01.06>

KOTHARI, C. R. *Research Methodology: Methods and Techniques*. 2. ed. New Delhi: New Age International (P) Limited, 2004. Acesso em 17 de jun de 2023. Disponível em: <<https://ccsuniversity.ac.in/bridge-library/pdf/Research-Methodology-CR-Kothari.pdf>>

LAVE, J. e WENGER, E. *Situated learning: Legitimate peripheral participation*. Cambridge University Press. 1991.

LERNER, Delia. *Ler e escrever na escola: o real, o possível e o necessário*. Porto Alegre: Artmed, 2002.

LEAL, S.E., DURAN, M.S., SCARTON, C., HARTMANN, N.S., e ALUÍSIO, S.M., "NILC-Matrix: assessing the complexity of written and spoken language in Brazilian Portuguese", arXiv e-prints, p. arXiv: 2201.03445, 2021. DOI:i:10.48550/arXiv.2201.03445.

LEAL, S.E., DURAN, M.S., SCARTON, C., HARTMANN, N.S., e ALUÍSIO, S.M. NILC-Matrix: Assessing the complexity of written and spoken language in Brazilian Portuguese. **Language Resources and Evaluation**, p. 1-38, 2023. Acesso em 15 de set de 2023. Disponível em: <<https://link.springer.com/article/10.1007/s10579-023-09693-w>> DOI: <https://doi.org/10.1007/s10579-023-09693-w>

LEE, C. J., SUGIMOTO, C. R., ZHANG, G. e CRONIN, B. Bias in peer review. **Journal of the American Society for Information Science and Technology**, 64(1), 2-17. 2013 DOI: <https://doi.org/10.1002/asi.22784>

LILLIS, T. e CURRY, M. J. Academic writing in a global context: The politics and practices of publishing in English. Routledge. **J Bus Tech Commun**, v. 22, p. 179-198, 2010. Disponível em: https://www.researchgate.net/profile/Theresa-Lillis/publication/42798576_Academic_Writing_in_a_Global_Context_The_Politics_and_Practices_of_Publishing_in_English/links/02bfe5114f531a02fc000000/Academic-Writing-in-a-Global-Context-The-Politics-and-Practices-of-Publishing-in-English.pdf

LINCK, A. C. B., SIGNORI, G. P., da SILVA, B. L., de ALMEIDA, C. A e BIGOLIN, M.. O RevisãoOnline: Plataforma on-line de apoio para a melhora da escrita de textos dissertativo-argumentativos. **Anais do Computer on the Beach**, v. 13, p. 341-346, 2022. DOI: [10.14210/cotb.v13.p341-346](https://doi.org/10.14210/cotb.v13.p341-346)

LIU, B. Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies. 2012. Acesso em 15 de dezembro de 2023. Disponível em: <<https://www.cs.uic.edu/~liub/FBS/SentimentAnalysis-and-OpinionMining.pdf>>

LIU, M., LI, Y., XU, W. E LIU, L., "Automated Essay Feedback Generation and Its Impact on Revision," in IEEE Transactions on Learning Technologies, vol. 10, no. 4, pp. 502-513, 1 Oct.-Dec. 2017, DOI: 10.1109/TLT.2016.2612659. Acesso em 16 mar 2023. Disponível em: <https://ieeexplore.ieee.org/abstract/document/7574351>

LONG, P., SIEMENS, G., "Penetrating the Fog: Analytics in Learning and Education, *Educause Review*, 46 (5), 31-40, 2011.

LOURENCO, Abílio A.; PAIVA, Maria Olímpia A. de. A motivação escolar e o processo de aprendizagem. *Ciênc. cogn.*, Rio de Janeiro, v. 15, n. 2, p. 132-141, ago. 2010. Disponível em <http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1806-58212010000200012&lng=pt&nrm=iso>. Acesso em: 05 mar. 2023.

LUCKIN R, HOLMES W, GRIFFITHS M, FORCIER LB. *Intelligence Unleashed: An argument for AI in Education*. London: Pearson Education; 2016. Disponível em: <https://discovery.ucl.ac.uk/id/eprint/1475756/1/Pearson%20Intelligence%20Unleashed%20FINAL.pdf>

MANNING, C. D. e SCHUTZE, H. "Foundations of Statistical Natural Language Processing". MIT Press. 1999.

MARCUSCHI, L. A. Gêneros textuais: definição e funcionalidade. In: DIONÍSIO, A. P.; MACHADO, A. R.; BEZERRA, M. A. (Orgs.). *Gêneros textuais e ensino*. Rio de Janeiro: Lucerna, 2010. p. 19-36. Disponível em: http://moodle.stoa.usp.br/file.php/752/bibliografia/marcuschi_2003.pdf. Acesso em: 12 de fev. de 2023.

MARTINS, Mário Gleisse das Chagas. *Complexidade textual e progressão escolar em dois registos : um estudo de correlação baseado em um corpus quasi-longitudinal*. 2016. Tese de doutorado, Universidade de Lisboa, 2016. Disponível em: <https://repositorio.ul.pt/handle/10451/23963>. Acesso em: 29 out. 2023.

MCNAMARA, Danielle S; GRAESSER, Arthur C. Coh-Metrix An Automated Tool for Theoretical and Applied Natural Language . In: MCCARTHY, P; BOONTHUM-DENECKE, C (org.). *IGI Global eBooks*. [S. l.: s. n.], 2012. p. 188–205. E-book. Disponível em: <https://doi.org/10.4018/978-1-60960-741-8.ch011>. Acesso em: 29 out. 2023

MEIRA, R. R.; WEIAND, A.; REATEGUI, E.; BIGOLIN, M.; MOTZ, R. A Analítica da Escrita para Identificação de Indicadores de Qualidade Textual. *Revista Novas Tecnologias na Educação*, Porto Alegre, v. 21, n. 2, p. 342–351, 2023. Disponível em: <<https://seer.ufrgs.br/index.php/renote/article/view/137756>>. Acesso em: 3 jan. 2024.

MUKAKA M. M. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J*. 2012 Sep;24(3):69-71. PMID: 23638278; PMCID: PMC3576830. Acesso em 06 de abr de 2024. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3576830/>

NADEAU, D. e SEKINE, S. A survey of named entity recognition and classification. *Linguistica e Investigationes*. 2007. <https://doi.org/10.1075/li.30.1.03nad> Acesso em 15 de dezembro de 2023. Disponível em: <<https://nlp.cs.nyu.edu/sekine/papers/li07.pdf>>

NIVRE, J., de MARNEFFE, M. C., GINTER, F., GOLDBERG, Y., HAJIC, J., MANNING, C. D. e ZEMAN, D. Universal Dependencies v1: A Multilingual Treebank Collection. In Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC 2016). 2016. Acesso em 05 de dez de 2023. Disponível em: <<https://nlp.stanford.edu/pubs/nivre2016ud.pdf>>

NCTE. (2013). NCTE position statement on machine scoring [Web log post] Apr 20. Retrieved Oct 10, 2018. Disponível em: [http://www2.ncte.org/statement/machine scoring](http://www2.ncte.org/statement/machine%20scoring)

OLIVE, Thierry. Writing and working memory: A summary of theories and of findings. In: **Writing**. Psychology Press, 2012. p. 125-140. Disponível em: [https://www.researchgate.net/publication/258861343 Writing and working memory A summary of theories and of findings](https://www.researchgate.net/publication/258861343_Writing_and_working_memory_A_summary_of_theories_and_of_findings). Acesso em <02 de Mar de 2023>

NICOL, D. J. e MACFARLANE-DICK, D. Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218. 2006. DOI: <https://doi.org/10.1080/03075070600572090>

PANG, B. e LEE, L. Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*: Vol. 2: No. 1–2, pp 1-135. 2008. <http://dx.doi.org/10.1561/15000000011>. Acesso em 15 de dezembro de 2023. Disponível em: <<https://www.cs.cornell.edu/home/llee/omsa/omsa.pdf>>

PEREIRA, Luísa Álvares. A formação de professores para o ensino da escrita. In: Sim-Sim, Inês (Org), A formação para o Ensino da Língua Portuguesa na Educação Pré-Escolar e no 1º Ciclo do Ensino Básico, *Cadernos de Formação de Professores*, nº 2, 2001, p. 35-49. Acesso em: 12 de fev. de 2023. Disponível em: http://www.casadaleitura.org/portalbeta/bo/abz_indices/000705_FP.pdf .

PERRENOUD, Philippe. *Construir as competências desde a escola*. Porto Alegre: Artmed, 1999.

PIRES, Carla; CAVACO, Afonso; VIGÁRIO, Marina. Towards the Definition of Linguistic Metrics for Evaluating Text Readability. *Journal of Quantitative Linguistics*, [s. l.], v. 24, n. 4, p. 319–349, 2017. Disponível em: <https://doi.org/10.1080/09296174.2017.1311448>. Acesso em: 29 out. 2023.

POLITZER-AHLES, Stephen. GIROLAMO, Teresa. GHALI, Samantha. Preliminary evidence of linguistic bias in academic reviewing, *Journal of English for Academic Purposes*. v. 47, 2020. <https://doi.org/10.1016/j.jeap.2020.100895>.

PRAÇA, Fabíola Silva Garcia. Metodologia da pesquisa científica: Organização estrutural e os desafios para redigir o trabalho de conclusão. *Revista Eletrônica Diálogos Acadêmicos*, [S. l.], v. 08, n. 1, p. 72–87, 2015. Disponível em: <http://www.uniesp.edu.br/fnsa/revista> .

RADFORD, A. Wu, J., Child, R., Luan, D., Amodei, D., Sutskever, I. Language Models are Unsupervised Multitask Learners. OpenAI, v. 1, n. 8, p. 9. 2019. Acesso em 23 de dezembro de 2023. Disponível em: <<https://d4mucfpxsywv.cloudfront.net/better-language-models/language-models.pdf>>

REATEGUI, E., KLEMMANN, M., EPSTEIN, D., e LORENZATTI, A. Sobek: A text mining tool for educational applications. In *Proceedings of the International Conference on Data Science (ICDATA)* (p. 1). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp). 2011.

REATEGUI, E., BIGOLIN, M., CARNIATO, M., dos SANTOS, R.A. Evaluating the Performance of SOBEK Text Mining Keyword Extraction Algorithm. In: Holzinger, A., Kieseberg, P., Tjoa, A.M., Weippl, E. (eds) *Machine Learning and Knowledge Extraction. CD-MAKE 2022. Lecture Notes in Computer Science*, vol 13480. 2022. Springer, Cham. https://doi.org/10.1007/978-3-031-14463-9_15

REID, J. *Teaching ESL Writing*. New Jersey: Lawrence Erlbaum Associates. 2016. Acesso em 20 de dez de 2023. Disponível em: <<https://repo.iainbatusangkar.ac.id/xmlui/handle/123456789/2594>>

RESNIK, David B. *The Ethics of Science: An Introduction*. London: Routledge, 1998. Acesso em 20 de dez de 2023. Disponível em: <<https://www.taylorfrancis.com/books/mono/10.4324/9780203979068/ethics-science-david-resnik>>

ROSALES F, GARCIA A, RODRIGUEZ S, PEDRAZA J. L, MENDEZ R. e NIETO M. M, "Detection of Plagiarism in Programming Assignments," in *IEEE Transactions on Education*, vol. 51, no. 2, pp. 174-183, May 2008, doi: 10.1109/TE.2007.906778. Acesso em 20 de dez de 2023. Disponível em: <<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4455467>>

RUDESTAM, K. E. e NEWTON, R. R. *Surviving your dissertation: A comprehensive guide to content and process*. Sage Publications. 2015.

Russell, Margo K., "A Comparison of Linguistic Features in the Academic Writing of Advanced English Language Learner and English First Language University Students" (2014). *Dissertations and Theses*. <https://doi.org/10.15760/etd.2022>

SALDAÑA, J. *The coding manual for qualitative researchers*. London: Sage. 2013.

SEO, K., TANG, J., ROLL, I. The impact of artificial intelligence on learner–instructor interaction in online learning. *Int J Educ Technol High Educ* 18, 54. 2021. <https://doi.org/10.1186/s41239-021-00292-9>

SIEMENS, G. e BAKER, R. S. Learning analytics and educational data mining: Towards communication and collaboration. In **Proceedings of the 2nd International Conference on Learning Analytics and Knowledge**. p. 252-254. ACM. 2012. DOI: <https://doi.org/10.1145/2330601.2330661>

SILVERMAN, D. Qualitative research. Sage Publications., v. 3, n. 3, p. 14-25, 2016.

SHERRI, M. D. e BURSTEIN, J. Handbook of Automated Essay Evaluation: Current Applications and New Directions. Routledge. 2013. Disponível em: <https://api.taylorfrancis.com/content/books/mono/download?identifierName=doi&identifierValue=10.4324/9780203122761&type=googlepdf>

SOUSA, Luciano Dias de, ALMEIDA, Flávio Aparecido de, BARD, Lucimere Aleixo, CANCELA, Lucas Borcard. Os desafios enfrentados pelos professores no processo de avaliação no ensino superior. *Revista de Gestão e Avaliação Educacional*, vol. 7, núm. 16, pp. 59-66, 2018. Universidade Federal de Santa Maria - UFSM. Acesso em 10 de mar de 2023. Disponível em: <https://www.redalyc.org/journal/4718/471857006005/html/>

SWALES, J. M. e FEAK, C. B. Academic writing for graduate students: Essential tasks and skills (3rd ed.). University of Michigan Press. 2012.

TAI, H.-C., LIN, W. C. e YANG, S. C. Exploring the Effects of Peer Review and Teachers' Corrective Feedback on EFL Students' Online Writing Performance. *Journal of Educational Computing Research*, 53(2), 284–309. 2015. DOI: <https://doi.org/10.1177/0735633115597490>. Acesso em: 05 de abr de 2023. Disponível em: <https://journals.sagepub.com/doi/full/10.1177/0735633115597490?journalCode=jeca>.

TASKIRAN, Ayşe e GÖKSEL, Nil. AUTOMATED FEEDBACK AND TEACHER FEEDBACK: WRITING ACHIEVEMENT IN LEARNING ENGLISH AS A FOREIGN LANGUAGE AT A DISTANCE. **Turkish Online Journal of Distance Education**, v. 23, n. 2, p. 120-139, 2022. DOI: <https://doi.org/10.17718/tojde.1096260>

TIMMS, M.J. Letting Artificial Intelligence in Education Out of the Box: Educational Cobots and Smart Classrooms. **Int J Artif Intell Educ** 26, 701–712, 2016. <https://doi.org/10.1007/s40593-016-0095-y>

TINKLER, P. e JACKSON, C. Examining the doctorate: Institutional policy and the PhD examination process in Britain. *Studies in Higher Education*, 29(2), 135-150, 2004

THI, N.K., NIKOLOV, M. How Teacher and Grammarly Feedback Complement One Another in Myanmar EFL Students' Writing. *Asia-Pacific Edu Res* 31, 767–779, 2022.
<https://doi.org/10.1007/s40299-021-00625-2>

URMILA, Baida Siti. ANALYSIS OF COHERENCE AND COHESION IN THE STUDENTS' WRITING TEXT. Orientador: Dr. Jarjani Usman. Tese (Doutorado) - Department of English Language Education, Fakultas Tarbiyah dan Keguruan, UNIVERSITAS ISLAM NEGERI AR-RANIRY BANDA ACEH - Singapura. 2021. <Disponível em:
<https://repository.ar-raniry.ac.id/id/eprint/19054/1/Siti%20Baida%20Urmila.%20140203256.%20FTK.%20PBI.%20085371937725.pdf>> Acesso em: 10 jul. 2023.

VALENTI, S., NERI, F. e CUCCHIARELLI, A. An overview of current research on automated essay grading. **Journal of Information Technology Education: Research**, 2(1), 319-330, 2003. Disponível em: <https://www.learntechlib.org/p/111481/>

VASWANI, Ashish et al. Attention is all you need. **Advances in neural information processing systems**, v. 30, 2017.

VERGARA, Sylvia Constant. Projetos e relatórios de pesquisa em administração. [s.l: s.n.]. Disponível em:
http://www.worldcat.org/title/projetos-e-relatorios-de-pesquisa-emadministracao/oclc/685253421&referer=brief_results.

VYGOTSKY, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

VYGOTSKY, L. S. *A formação social da mente*. São Paulo: Martins Fontes, 1991.

VYGOTSKY, L. S. *A formação social da mente: o desenvolvimento dos processos psicológicos superiores*. São Paulo: Martins Fontes, 1998.

WANG, Z.; HAN, F. The Effects of Teacher Feedback and Automated Feedback on Cognitive and Psychological Aspects of Foreign Language Writing: A Mixed-Methods Research. *Frontiers in Psychology*, v. 13, 2022. Disponível em:
<https://www.frontiersin.org/articles/10.3389/fpsyg.2022.909802> . Acesso em: 12 fev. 2023.

WARSCHAUER, M. e WARE, P. Automated writing evaluation: defining the classroom research agenda. *Language Teaching Research*, 10(2), 157-180. 2006. DOI:
<https://doi.org/10.1191/1362168806lr190oa> Acesso em 18 de dez de 2023.

WARSCHAUER, M. New tools for teaching writing. **Language Learning & Technology**, (14)1, 3–8. 2010. DOI: <http://dx.doi.org/10125/44196> Acesso em 18 de dez de 2023.

WESTON-SEMENTELLI, Jennifer L.; ALLEN, Laura K.; MCNAMARA, Danielle S. Comprehension and Writing Strategy Training Improves Performance on Content-Specific Source-Based Writing Tasks. *International Journal of Artificial Intelligence in Education*, [S. l.], v. 28, n. 1, p. 106–137, 2018. DOI: 10.1007/s40593-016-0127-7. Acesso em 18 de dez de 2023.

WILSON, Joshua; OLINGHOUSE, Natalie G.; ANDRADA, Gilbert N. **Does Automated Feedback Improve Writing Quality?** *Learning Disabilities: A Contemporary Journal*, v12 n1 p93-118 Mar 2014 Disponível em: <<https://files.eric.ed.gov/fulltext/EJ1039856.pdf>> Acesso em 18 de dez de 2023.

WOODWORTH Johanathan, HYBRID FEEDBACK: THE EFFICACY OF COMBINING AUTOMATED AND TEACHER FEEDBACK FOR SECOND LANGUAGE ACADEMIC WRITING DEVELOPMENT. [Doctoral dissertation, *York University*]. *YorkSpace*. 2022 DOI: <https://yorkspace.library.yorku.ca/xmlui/handle/10315/39657> Acesso em 18 de dez de 2023.

XIA, Menglin; EKATERINA KOCHMAR; BRISCOE, Ted. Text Readability Assessment for Second Language Learners. In: BEA@ACL, 2016, San Diego, CA. 11th Workshop on Innovative Use of NLP for Building Educational Applications. [S. l.]: Association for Computational Linguistics, 2016. Disponível em: <<https://aclanthology.org/W16-0502.pdf>> Acesso em: 15 out. 2023.

YIN, Robert K. *Case Study Research: Design and Methods*. 5. ed. Thousand Oaks: SAGE Publications, 2014. Acesso em 20 de jul de 2023. Disponível em: <https://www.researchgate.net/publication/308385754_Robert_K_Yin_2014_Case_Study_Research_Design_and_Methods_5th_ed_Thousand_Oaks_CA_Sage_282_pages>

ZHAI, N., MA, X. The Effectiveness of Automated Writing Evaluation on Writing Quality: A Meta-Analysis. *Journal of Educational Computing Research*, 0(0):1-26, 2022. DOI: 10.1177/07356331221127300 Acesso em 18 de dez de 2023.

ZHANG, Shuai. "Review of automated writing evaluation systems" *Journal of China Computer-Assisted Language Learning*, vol. 1, no. 1, 2021, pp. 170-176. <https://doi.org/10.1515/jccall-2021-2007>. Acesso em 18 de dez de 2023. Disponível em: <<https://www.degruyter.com/document/doi/10.1515/jccall-2021-2007/html>>