DEBISS：个人、半结构化口语辩论语料库*

Klaywert Danillo Ferreira de Souza[1],**David Eduardo Pereira**[1],
Cl´音频 E. C. Campelo[1],**Larissa Lucena Vasconcelos**[2]

[1]坎皮纳格兰德联邦大学（UFCG）- 系统与计算系 坎皮纳格兰德，巴西

[2]帕拉伊巴联邦学院（IFPB）蒙泰罗，
巴西

klaywertdanillo@copin.ufcg.edu.br , david.pereira@ccc.ufcg.edu.br

campelo@dsc.ufcg.edu.br，larissalucena@gmail.com

arXiv:2603.05459v1 [cs.CL] 5 Mar 2026

摘要。 辩论在我们的日常生活中至关重要，无论是在学习、工作、日常讨论、电视上的政治辩论，还是社交网络上的在线讨论中都是如此。辩论的应用范围十分广泛。由于辩论的应用场景、结构和形式各不相同，要开发能够涵盖这些差异的语料库并非易事，而现有研究中的辩论语料库也相当匮乏。因此，本研究提出 *DEBISS* 语料库——一个包含口语个人辩论并具有半结构化特征的语料库。该语料库提供丰富的自然语言处理任务标注，例如语音转文本、说话人分离、论证挖掘和辩手质量评估。

1. 引言

辩论在教育中发挥着关键作用，能够显著提升批判性思维和沟通能力。尽管辩论十分重要，但由于缺乏多样且标注丰富的语料库，辩论话语的计算分析进展受到了限制。这一局限制约了研究方向的拓展，也阻碍了针对细致复杂的辩论互动开发专门计算模型。

此外，教育环境中口语、半结构化辩论的资源尤其匮乏，英语以外的语言更是如此。为填补这一空白，我们推出了 DEBISS（口语、个人及半结构化辩论），这是一个全新的、经过标注的巴西葡萄牙语语料库。该语料库收录了总计 9 小时 35 分钟的面对面、个人、半结构化辩论，主题围绕“生成式人工智能及其对社会的影响”，参与者为来自坎皮纳格兰德联邦大学的 67 名计算机科学专业一年级学生。

因此，DEBISS语料库提供了一个独特的数据集，聚焦于口语形式的个人辩论，其中融合了预设问题和自由表达的环节，使其能够呈现出不同于高度结构化或书面形式的自发论证。

*本文的缩略版已发表于 STIL 2025 论文集。如果您希望引用我们的研究，请参阅已发表的版本：https://doi.org/10.5753/stil.2025.37860。您也可以引用以下相关研究： [https://doi.org/10.5753/stil.2025. 37836，https://doi.org/10.5753/jbcs.2025.5824]。

其主要贡献包括聚焦巴西葡萄牙语，丰富辩论分析中的语言多样性；立足于教育场景，为了解学生演讲技能发展提供了见解。此外，该语料库还收录了参与者针对自身表现和主题知识所做的详细自我评估与同伴评估，以及对全面分析至关重要的多模态数据（音频和转录文本）；其主题紧跟时代，也确保了内容的现实意义。

这个综合性语料库有助于探索论辩话语的各个方面，从论证构建和修辞策略到辩手互动模式；它对于自然语言处理（NLP）任务也很有价值，例如论证挖掘（AM）、辩手质量分析（DQA）、语音转文本、说话人分离、言语不流畅检测等。DEBISS语料库（包括所有转录文本、音频文件及相关标注）将通过GitHub[1]向研究用途开放。

本文其余部分的结构如下：第2节回顾与辩论语料库相关的研究，并考察该领域数据集构建所采用的知识表示和方法论。第3节详细介绍数据收集与处理、语料库分析及统计见解的方法论。最后，第4节将介绍有关该数据集的结论。

2. 相关工作

辩论的研究与分析在多个领域都受到了广泛关注，包括计算语言学、话语分析和人工智能。研究人员开发并使用了众多语料库，以理解和建模论辩话语的不同方面。本节回顾重要的辩论语料库及其方法论，重点关注其结构、标注方案，以及它们在论证挖掘（AM）和辩手质量分析（DQA）等计算任务中的适用性。

2.1. 政治辩论语料库

由于政治辩论结构严谨且公开可得，因此一直是构建论辩话语语料库的主要研究对象。其中较为著名的是美国总统辩论语料库 [Vrana 和 Schneider 2017], 以及其他政治来源[Duthie 等，2016；Mestre 等，2021；Mancini 等， 2022], ，该语料库包含多场总统辩论的文字记录，并对其中的论证结构和修辞策略进行了标注 [De Smedt 和 Jaki，2018；Carvalho 等，2011；Hautli-Janisz 等， 2022]。尽管政治辩论语料库颇有用处，但其正式程度通常较高，且遵循严格规范，未必能反映结构较松散或自发性较强的论辩互动。此外，由于聚焦于特定情境和主题，基于这些数据集训练的模型难以推广到其他形式的辩论和话语中。

2.2. 在线辩论与讨论语料库

在线平台的兴起促成了基于数字讨论和辩论构建的语料库 [Durmusand Cardie2019, Khodaketal. 2017,

[1]https://github.com/AINDA-Project-UFCG/DEBISS

Stranisci et al. 2021, Lai et al. 2018]。关于辩论的研究涉及在线平台上的书面辩论 [Sousa et al. 2021, Habernal and Gurevych 2016, ˇBoltuˇzi´c and Snajder 2016, Chakrabarty et al. 2019], , 例如 Twitter 和红迪。

互联网论证语料库（IAC） [Abbott 等人 2016] 汇集了来自红迪和论坛等平台的讨论，并对论证立场、质量和相关性进行了标注。该语料库涵盖了广泛的主题和风格，为研究非正式论辩话语和情感分析提供了丰富资源。在线辩论语料库常面临噪声、非正式语言和话语结构各异等问题，给一致的标注与分析带来挑战。此外，这些数据集主要呈现书面交流，缺少口头辩论中存在的多模态和自发性元素。

2.3. 教育与学术辩论语料库

学术环境为研究论辩技能发展和话语能力提供了沃土。大多数教育语料库 [Ruiz-Dolz 等人 2021] 收录了学术正式辩论的录音和文字记录，揭示了教育话语的特点，但并未特别关注结构较为松散的辩论。现有教育语料库通常并未明确聚焦辩论，尤其是半结构化且包含自发论证的辩论。经过标注、能够捕捉学生主导辩论细微特点的数据集十分匮乏，尤其是英语以外语言的数据集，这限制了论辩话语分析领域的跨文化与多语言研究。

3. 方法论

辩论是日常生活中经常发生的一种人际互动形式，并随着时间的推移演变为交流的基本方面。本研究旨在构建一个特征明确的辩论语料库。为实现这一目标，我们制定了数据收集与处理的方法论，并在整个语料库创建过程中遵循特定准则。本节详细说明了在拟议语料库开发过程中设计并实施的方法步骤。

3.1. DEBISS语料库

DEBISS语料库包含经（已省略）大学计算机科学专业一年级本科生同意后开展的现场辩论的音频转录文本。每场辩论均由一名主持人主持。数据收集于2024年进行，共有**67** 名学生参与，他们被分成**16**个辩论小组，共计生成**9小时** 和**35分钟** 的录音。这些录音通过结合语音转文本人工智能模型与人工验证的半自动化流程进行转录，以确保准确性。

此外，不同语料库的辩论形式可能有所不同：有些采用独白式风格，有些由个人辩手捍卫个人观点，还有一些采用小组式方法，由集体共同捍卫单一立场。在本研究中，辩论小组由**3** 至**5**名参与者组成，共计16个小组。每位参与者独立发言，捍卫自己的观点。这种以个人为中心的形式与其他语料库所采用的方法一致，这些方法强调每位辩论者的个人论辩特征。

3.1.1. 辩论主题

为了开展一场聚焦且引人入胜的辩论，选定了一个中心主题：“生成式人工智能及其对社会的影响”。该主题高度相关且具有争议性，引发了各领域广泛的质疑与批评。在一些语料库方法中，确定中心主题以引导辩论是常见做法。然而，一些语料库，尤其是聚焦政治辩论的语料库，并不局限于特定主题，因为政治格局的多个方面——如安全、国际关系、医疗保健和教育——都可以加以讨论。

此外，为了推动辩论进程，研究团队汇编了一组线上文本，供辩论者在辩论前阅读。这些文本因易于阅读而入选，优先选择可靠来源的新闻文章和评论文章，而非复杂的学术论文。所选文本围绕主题，探讨了具体、相关且有争议的问题，例如人工智能对工作岗位的影响、生成式人工智能在教育中的应用，以及人工智能解决方案的法律影响。由于阅读所选文本并非强制要求，研究团队还提供了一份两页的摘要，称为“补充文本”。该摘要提炼了关键议题的核心信息。

3.1.2. 辩论环境设置

为从辩论中收集详细数据，研究人员使用处于全向模式的罗技 USB Yeti 电容麦克风，并配合 OBS Studio 软件录制各场辩论环节。音频在一间 3×5 米的会议室中采集，辩论者围坐在一张中央桌子旁。OBS 的设置经过调整，以确保所有参与者的分贝水平达到最佳，录制生成的 MKV 文件随后被转换为 MP3 格式。房间内还配有一台显示辩论信息的 55 英寸电视，布局如随附的图3所示。

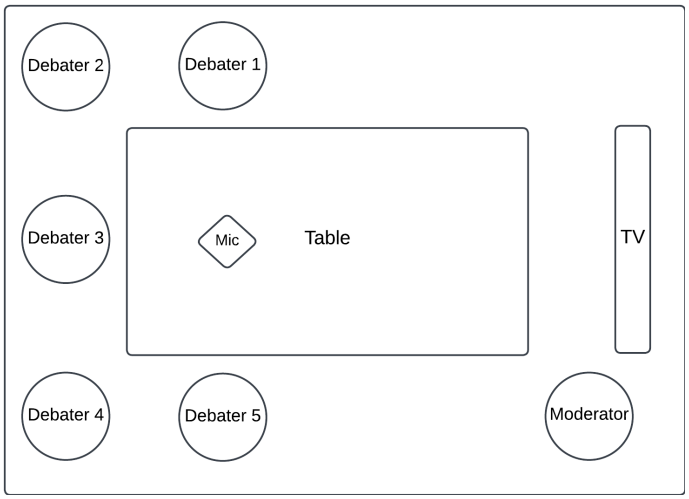


图 1. 辩论环节期间的房间布局

3.1.3. 辩论形式

数据收集方案旨在通过半结构化辩论采集语音数据。所有辩论环节均遵循标准化的半结构化形式，以确保数据收集过程始终保持一致。环节开始时会说明研究目的，并告知自愿参与者，辩论将被录制，但不会包含任何个人信息。主要目标是建立一个经过转录的辩论语料库，供各种场景使用，尤其是用于开发专注于辩论分析的人工智能模型。以下将详细介绍这些环节的结构：

1. 介绍并说明创建语料库的动机 2. 签署声音和图像使用同意书 3. 为每位辩论者分配唯一的数字标识符 4. 录制辩论者的语音样本，以便识别身份 5. 说明辩论规则 6. 陈述对议题的初始立场和想法 7. 问答环节 8. 陈述对议题的最终立场并分享感想 9. 结束录音 10. 填写自我评估表

接下来，请参与者签署同意书，授权将其录制的声音用于研究用途。这一步至关重要；由于研究涉及人类受试者，研究已正式获得该大学伦理委员会的批准。为便于人工分析时识别声音，每位辩论者都会在辩论开始前先录制同一句话。收集到签署的同意书和语音样本后，辩论环节便正式开始。主持人负责监督辩论，维持秩序并鼓励讨论。主持人首先说明规则：辩论者应听从主持人，不要互相打断；如需发言，应举手示意，并等待轮到自己发言，或等当前发言者说完。

此外，辩论环节分为三个部分。**第一部分** 允许辩论者表达他们对主题的初步意见。**第二部分** 是问答环节，**第三部分** 则围绕主题分享最终想法并进行反思。每位辩论者都会收到指定问题，由主持人朗读，并显示在电视上。回答完毕后，其他辩论者可以发表评论或提出补充问题。这种互动是自愿的，为交流提供了开放空间，同时保持半结构化形式，兼有强制互动和自愿互动。每位辩论者都有不受打断的时间来表达意见和回答问题。共准备了五个问题，每位辩论者一个；如果只有三名参与者，则只使用前三个问题。

问答环节结束后，会向所有参与者提出最后一个问题，让有意愿的人补充更多见解。此互动同样是自愿的，营造出丰富且富有参与性的辩论氛围，鼓励交流多样的观点与见解。在辩论的最后部分，参与者需要分享任何总结性想法，并说明自己的观点是否因讨论而有所改变。这项最终反思是必答的，让每位辩手都有机会表达自己的最终立场。

3.2. 辩手评估

辩论和录制结束后，每位参与者都填写了一份自我评估表（Google 表单）。这对于评估个人表现以及对辩论过程的个人反思至关重要，也为分析提供了宝贵的数据。这些问题旨在了解辩手的准备情况、参与程度和整体体验。识别编号确保回复能够与辩论数据关联，同时维护保密性。自我评估表包含以下问题：

- 您的编号是什么？（多项选择）
- 您如何评价自己在辩论中的表现？（李克特量表：差至优秀）
- 辩论开始前，您对辩题的了解程度如何（通过学习或经验）？（李克特量表：非常不同意至非常同意）
- 您是否仔细阅读了组织者提供的材料？（李克特量表：非常不同意至非常同意）
- 您是否学习了所提供材料以外的补充材料？（李克特量表：非常不同意至非常同意）
- 这场辩论是否拓展了你对该主题的视角或知识？(李克特量表：非常不同意至非常同意)
- 你认为谁是表现最佳的辩论者？（单项选择，包括“没有最佳辩论者”选项）
- 请说明你选择这位最佳辩论者的理由。（开放式文本）

参与者对自己的表现进行评分，并反思自己在论证中的有效性，从而让人们了解他们在辩论动态中的自我认知。关于先前知识和准备情况的问题，评估了参与者对主题的熟悉程度以及他们对所提供材料的投入情况。这有助于评估准备工作如何影响其表现。关于辩论是否拓宽了参与者视角的问题，旨在衡量这场活动的教育价值。

此外，辩手评估还包括同伴评估：参与者选出小组中表现最佳的辩论者并说明理由，从多个角度揭示了群体动态和个人表现。总体而言，自我评估表全面收集了有关辩论准备情况、表现和教育影响的反馈。这些数据对于分析辩论结果、了解促成成功辩论的因素具有重要价值。此外，这对参与者的发展至关重要，有助于培养元认知能力并促进自我提升。通过分析这些数据，研究人员可以找出优秀辩论者所具备的关键技能和策略，为提升学术和专业情境中的辩论能力提供理论见解和实践指导。

3.3. 数据处理与标注

录制完所有辩论环节后，我们还需要进行自动和人工后处理。本节将详细介绍所采用的流程。第一步是转写录制音频文件，这是一项耗时的工作。为加快这一过程，我们使用免费模型进行自动转写，并测试了三种机器学习模型：wav2vec-large [2]，

[2]https://huggingface.co/facebook/wav2vec2-large

whisper-large [3] 和 Azure 语音转文本 [4]。比较它们在相同录音上的表现后，我们发现 Azure 的模型最准确，也最适合我们的使用场景。随后，我们使用 Azure 转写了所有音频文件，并将结果导出为 CSV 文件，其中包含以下信息：

- **转录文本**：从音频片段转录出的文本。
- **开始时间**：转录音频片段的开始时间。
- **结束时间**：转录音频片段的结束时间。

我们对自动转写中的错误进行了人工修正。最初的自动转写提供了一个有用的起点，大幅减少了人工工作量。我们使用 Google 表格作为标注工具，因为它便于编辑文本和时间戳。标注人员遵循特定流程来修正转录文本中的错误。首要目标是通过聆听音频、将其与自动转写进行对比并作出必要修正，确保准确性。另一项任务是处理多名辩论者同时发言的情况。尽管由于辩论规则，这类情况很少发生，标注人员仍会将混合发言拆分为多行，每行对应一名发言者。

此外，我们还进行了最终标注，以识别音频中的发言者。我们在电子表格中新增一列，为每段发言标注相应辩论者的标识符。另外，如果自动转写因停顿而将发言拆分成较小片段，标注人员会将这些片段合并为一行，并相应调整开始和结束时间。图 2 展示了该数据集的完整标注流程。

3.3.1. DEBISS 语料库统计

所提出的语料库旨在通过引入一个新的录音数据集，推动该领域的发展；该数据集有望为辩论的科学研究作出重要贡献。DEBISS 的独特之处在于，它纳入了区别于现有辩论语料库的鲜明特征，包括半结构化格式、个人发言者的贡献以及口语话语。该语料库旨在填补有关非典型辩论结构的文献空白，并支持广泛的自然语言处理应用。表 1 汇总了应用完整方法论后得到的语料库统计数据。

3.3.2. DEBISS 适用性

作为展示 DEBISS 语料库适用性的案例，我们重点介绍基于所收集数据构建的两个应用案例，分别名为 DEBISS-Arg 和 DEBISS-Eval。

[3]https://huggingface.co/openai/whisper-large-v3[4] https://learn.microsoft.com/en-us/azure/ai-services/speech-service/ speech-to-text[5]文本中的词汇变化度量

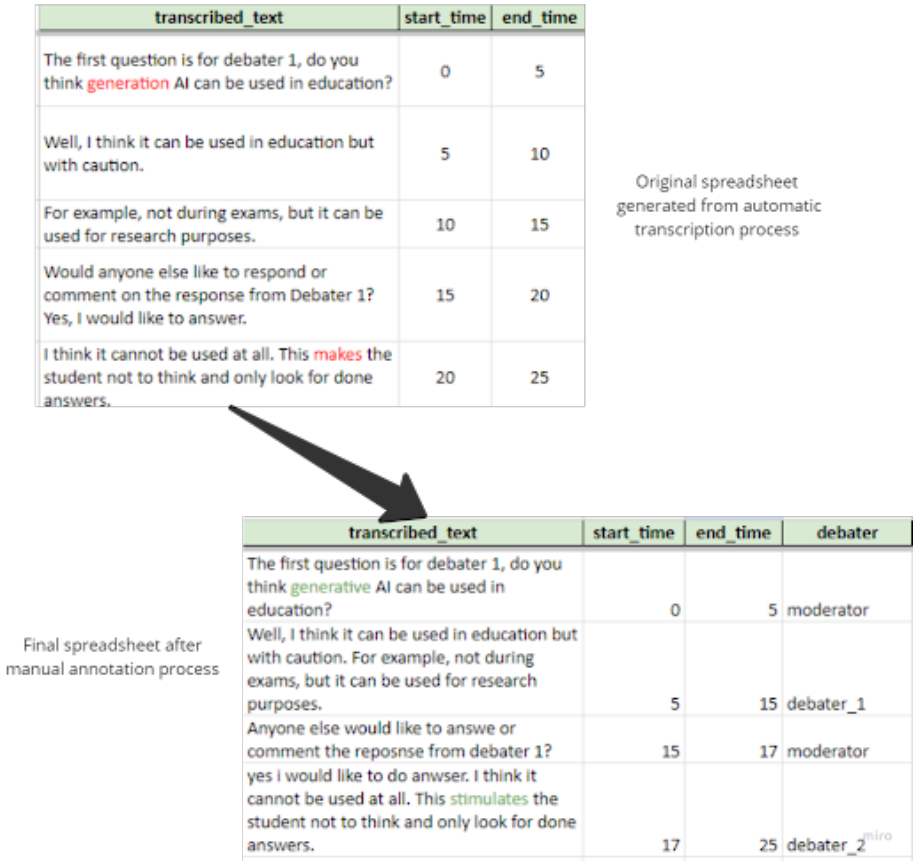


图 2。标注流程示例

表 1。DEBISS 统计数据

统计项	值
辩论者数量	67
小组数量	16
录音总时长	9小时37分钟
词元数量	130697
词汇多样性[5]	0.062

DEBISS-Arg 基于本研究介绍的新型 DEBISS 语料库，我们开发了一个名为 DEBISS-Arg 的衍生语料库，以推动论证挖掘领域的发展。该语料库经过全面标注，旨在支持各类论证挖掘任务的评估。具体而言，它包含对论证话语单元（ADUs）和非论证话语单元文本片段的标注，也标注了前提、主张和证据等论证组成部分。此外，该语料库还提供微观层面的关系标注，用于捕捉单个论证话语单元内部各组成部分之间的互动；以及宏观层面的关系标注，用于描绘辩论中不同话语发言之间的关系，从而捕捉辩论者之间的论证互动。

这套全面的标签使 DEBISS-Arg 语料库对于各种论证挖掘任务都极具价值。在创建这一资源的过程中，严格遵循了数据标注和处理方法论。完整的方法论细节，包括

从原始 DEBISS 语料库构建 DEBISS-Arg 所采取的步骤，详见 [Pereira 等人 2025] 。图 3 展示了一个示例，说明上述标签如何应用于 DEBISS-Arg 标注框架。

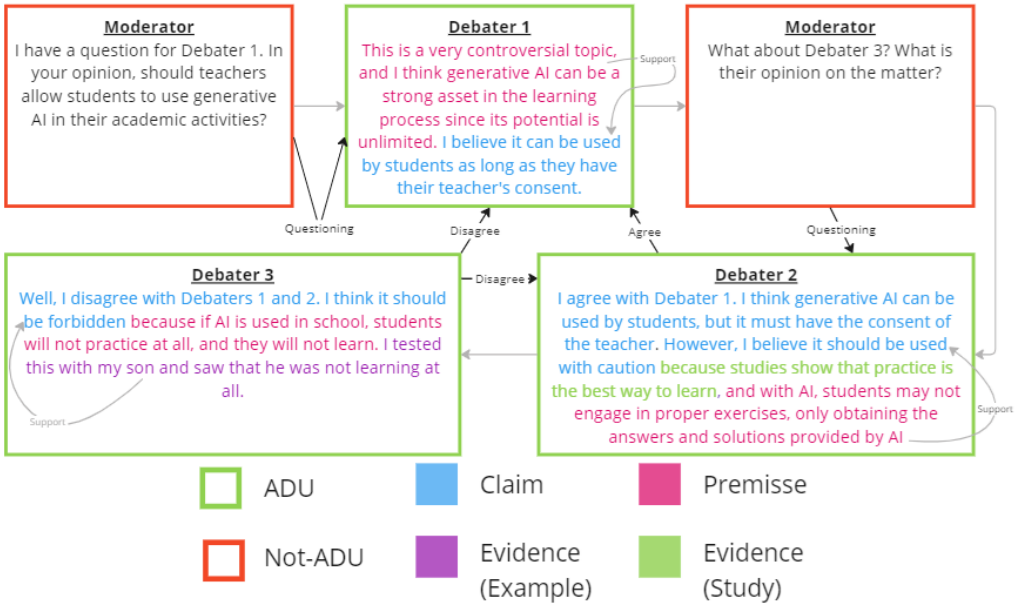


图3。DEBISS-Arg 数据示例（摘自 [省略]）

DEBISS-Eval DEBISS-Eval[6], 作为 DEBISS学术辩论集的一个专门子语料库，旨在推动对辩论质量和个体辩手表现评估的研究。由五位专家评委组成的专门小组，成员均拥有语言学学位且具备辩论经验，用两周时间细致评估了全部 16 场辩论。他们依据组织能力、论证、说服力和清晰度等既定标准，采用 1 至 5 分的李克特量表对辩论者进行评估，形成了丰富的数据，包括定量评分、硬投票和软投票结果，以及大量定性反馈 [Ericson 2011]。

此外，定性数据还包括评委对几乎所有辩手发言所做的自然语言笔记，以及对其整体“最佳辩手”评选结果的详细理由说明，最终汇集成 80 份详尽的评估反馈。评委笔记示例见图4。这些丰富的信息为深入分析论证有效性、辩论动态，以及论辩情境中人类评价判断的复杂性奠定了坚实基础。

不流利现象检测 DEBISS数据集也成为文本不流利现象检测研究的重点 [de Lima 和 Campelo 2024]。研究人员探索了大语言模型（LLMs）使用 DEBISS 完成这项任务的表现。该研究专门评估了 OpenAI 的 GPT-4o 等先进大语言模型 [OpenAI 等人 2024] ，考察它们能否从转录语音中识别并删除重复、犹豫和修正等不流利成分。研究采用了多种提示工程技术——零样本、少样本和思维链提示——对其进行全面测试

[6]可在 https://github.com/AINDA-Project-UFCG/DEBISS-Eval 获取

transcription	start_time	end_time	debater	judge1_notes	judge2_notes	judge3_notes	judge4_notes	judge5_notes
Bom, as IAs generativas são bem dizer uma faca de dois gumes, né? Tem aspectos positivos e muito negativos também, positivos tem automação automação de algumas atividades é pode auxiliar como ferramenta complementar em estudos como negativa, tem a questão de proteção de dados, de autenticidade, de resposta, infração de direitos autorais. Então, dependendo da área do argumento que quiser seguir, pode ser positivo ou negativo.	278.86	313.58	DEBATER1	Opinião geral do tema, comum a inicios de debates. Sem aprofundamentos, mas demonstrando hadaptabilidade relacionada a pontos positivos e negativos das inteligências artificiais generativas	Opinião geral, objetiva e direta. Cumprinido com o que foi solicitado por parte do moderador.	O debatedor me pareceu não estar muito preparado para responder, de modo que ele expos sua opinião espontaneamente. E por isso, acaba se expressando em poucas palavras.		Debatedor busca não se posicionar perante a temática e fornece apenas uma ideia geral sobre o tema, conforme solicitado.

图4：DEBISS-Eval 中评委对某一论证的评语示例

这些大语言模型使用 DEBISS数据集的能力。尽管一些模型表现有限，GPT-4o 在消除言语不流畅现象并保持文本质量方面始终表现出色，往往优于其他模型。这些发现凸显了 DEBISS 作为宝贵资源的重要性，它有助于开发和评估自动言语不流畅检测领域的前沿解决方案。

4. 结论

本研究构建了所提出的语料库，旨在通过引入一种新的录音语料库，推动辩论分析这一科学领域的发展，提升该领域的研究水平。DEBISS 还融入了若干鲜明特征，使其有别于现有的辩论语料库，例如半结构化特征、个人辩论和口语内容。此外，本研究的贡献不止于数据，还提出了一种按指定格式收集辩论内容的方法论

此外，该语料库旨在纳入丰富的标注，以支持多种自然语言处理任务。这些标注涵盖语音转文本转录、声纹识别、说话人分离、静音检测、不流利现象检测、辩手质量评估（DEBISS-Eval）以及论证挖掘任务（DEBISS-Arg）。因此，该语料库显著拓展了巴西葡萄牙语自然语言处理研究可用的开放数据。

尽管该数据集具有明确的方法论和详尽的统计数据，但仍需指出其中一些重要局限。其中一项局限是辩论的主题范围较窄，这可能会影响数据集的普适性。扩大辩论主题的范围并纳入更多元的主题内容，将有助于提升其适用性。此外，从背景和教育水平各异的学生等更多样化的参与者群体中收集数据，也将进一步增强数据集的代表性。未来的研究可将所提出的方法论应用于更广泛的主题以及更多元的辩手概况，从而解决这些局限。

参考文献

[Abbott 等人 2016] Abbott, R., Ecker, B., Anand, P., and Walker, M. (2016). 互联网论证语料库 2.0：面向对话式社交媒体的 SQL架构及配套语料库。载于 第十届国际语言资源与评估会议论文集（*LREC'16*），第 4445–4452 页。

ˇ ˇ[Boltuˇzi´c 和 Snajder 2016] Boltuˇzi´c, F. and Snajder, J. (2016). 填补空白！分析在线辩论中主张之间的隐含前提。载于 Reed, C.（编），第三届论证挖掘研讨会论文集（*ArgMining2016*），第 124–133 页，德国柏林。计算语言学协会。

[Carvalho 等人 2011] Carvalho, P., Sarmento, L., Teixeira, J., and Silva, M. J. (2011). 政治辩论评论情感标注语料库中的说谎者与救世主。载于计算语言学协会第*49*届年会论文集：人类语言技术，第 564–568 页。

[Chakrabarty 等人 2019] Chakrabarty, T., Hidey, C., Muresan, S., McKeown, K., and Hwang, A. (2019). AMPERSAND：面向说服性在线讨论的论证挖掘。载于 Inui, K., Jiang, J., Ng, V., and Wan, X. 编，*2019* 年自然语言处理实证方法会议与第九届自然语言处理国际联合会议（*EMNLP-IJCNLP*）论文集，第 2933–2943 页，中国香港。计算语言学协会。

[de Lima 和 Campelo 2024] de Lima, P. L. and Campelo, C. E. (2024年). 基于大语言模型的语音转写文本中不流畅表达的检测与去除。载于第十五届巴西人类语言信息技术研讨会论文集´¸ ˜，第 227–235 页，阿雷格里港，南里奥格兰德州，巴西。SBC。

[De Smedt 和 Jaki 2018] De Smedt, T. and Jaki, S. (2018). Polly语料库：德国在线政治辩论。载于第六届计算机媒介传播（*CMC*）与社交媒体语料库会议（*CMC-corpora 2018*），第 33 页。

[Durmus 和 Cardie 2019] Durmus, E. and Cardie, C. (2019). 一个用于建模在线辩论中论证的用户与语言效应的语料库。*arXiv*预印本 *arXiv:1906.11310*。

[Duthie 等人 2016] Duthie, R., Budzynska, K., and Reed, C. (2016). 在政治辩论中挖掘品格诉求，见《人工智能与应用前沿》第287卷，第299–310页。IOS Press，荷兰。本研究部分由英国工程与自然科学研究委员会（EPSRC）通过资助项目 EP/M506497/1 提供支持，部分由波兰国家科学中心通过资助项目 2015/18/M/HS1/00620 提供支持。

[Ericson 2011] Ericson, J. (2011).《辩论者指南》。

[Habernal 和 Gurevych 2016] Habernal, I. and Gurevych, I. (2016). 哪种论证更有说服力？使用双向长短期记忆网络分析和预测网络论证的说服力。载于 Erk, K. 和 Smith, N. A. 编，《第*54*届计算语言学协会年会论文集（第*1*卷：长篇论文）》，第1589–1599页，德国柏林。计算语言学协会。

[Hautli-Janisz 等人 2022] Hautli-Janisz、A.、Kikteva、Z.、Siskou、W.、Gorska、K.、Becker、R. 和 Reed、C.（2022）。Qt30：广播辩论中的论证与冲突语料库。载于第*13*届语言资源与评测会议论文集，第3291–3300页。欧洲语言资源协会（ELRA）。

[Khodak 等人 2017] Khodak, M., Saunshi, N., and Vodrahalli, K. (2017). 用于讽刺的大型自标注语料库。*arXiv* 预印本 *arXiv:1704.05579*。

[Lai 等人 2018] Lai, M., Patti, V., Ruffo, G., and Rosso, P. (2018). 意大利政治辩论中的立场演变与 Twitter 互动。载于自然语言处理与信息系统：第*23*届自然语言应用于信息系统国际会议，*NLDB 2018*，法国巴黎，*2018*年*6*月*13*日至*15*日，论文集*23*，第15–27页。Springer。[Mancini 等人 2022] Mancini, E., Ruggeri, F., Galassi, A., and Torroni, P. (2022). 多模态论证挖掘：政治辩论案例研究。载于 Lapesa, G., Schneider, J., Jo, Y., and Saha, S. 编，第*9*届论证挖掘研讨会论文集，第158–170页，线上举办及大韩民国庆州。计算语言学国际会议。[Mestre 等人 2021] Mestre, R., Milicin, R., Middleton, S. E., Ryan, M., Zhu, J., and Norman, T. J. (2021). M-arg：用于政治辩论的多模态论证挖掘数据集，包含音频和文字记录。载于 Al-Khatib, K., Hou, Y., and Stede, M. 编，第*8*届论证挖掘研讨会论文集，第78–88页，多米尼加共和国蓬塔卡纳。计算语言学协会。[OpenAI 等人 2024] OpenAI, :, Hurst, A., Lerer, A., Goucher, A. P., Perelman, A., and et. al, A. R. (2024). GPT-4o 系统卡片。[Pereira 等人 2025] Pereira, D., Sim˜ao, D., and Campelo, C. (2025). DEBISS-Arg：半结构化辩论中论证挖掘的深入数据标注协议与语料库。载于巴西信息技术与人类语言研讨会论文集 ´巴西信息技术¸ ˜与人类语言研讨会，第334–348页，巴西南里奥格兰德州阿雷格里港。SBC。[Ruiz-Dolz 等人 2021] Ruiz-Dolz, R., Nofre, M., Taul´e, M., Heras, S., and Garc´ıa-Fornes, A. (2021). Vivesdebate：一个全新的辩论赛论证多语标注语料库。《应用科学》，11(15):7160。[Sousa 等人 2021] Sousa, J. P., Nascimento, R., Araujo, R., and Coelho, O. (2021). 别在辩论中迷失！社交网络中的论证挖掘。载于巴西社交网络分析与挖掘研讨会论文集，第139–150页，巴西南里奥格兰德州阿雷格里港。SBC。[Stranisci 等人 2021] Stranisci, M., De Leonardis, M., Bosco, C., and Patti, V. (2021). Twitter 辩论中道德价值观的表达：对话语料库。*IJCoL*。《意大利计算语言学期刊》，7(7-1, 2):113–132。[Vrana 和 Schneider 2017] Vrana, L. and Schneider, G. (2017). 不择手段：创建并分析美国总统辩论文字记录语料库。