

Uma Ferramenta para Organização Semântica de Informações de Helpdesk por meio de Modelagem de Tópicos

Daniel Carvalho¹, Antônio Pereira¹, Washington Cunha², Elisa Tuler¹, Diego Dias³, Leonardo Rocha¹

¹Departamento de Ciência da Computação – Universidade Federal de São João Del-Rei (UFSJ), ²Departamento de Ciência da Computação – Universidade Federal de Minas Gerais (UFMG), ³Departamento de Estatística – Universidade Federal do Espírito Santo (UFES)

daniel@daniellima.net;antonio258p@gmail.com;washingtoncunha@dcc.ufmg.br
diego.dias@ufes.br;etuler@ufs.edu.br;lcrocha@ufs.edu.br

ABSTRACT

In this work, we propose a semantic organization tool for textual data applied to enhance the automated analysis of *helpdesk tickets* (service requests) in large volumes of data. It leverages Topic Modeling (TM) techniques (i.e., CluWords) and Large Language Models (LLMs) to identify recurring patterns and themes (Llama3), aiming to improve categorization and service efficiency. Through a case study with more than 9,000 *tickets*, we demonstrate its application in real-world scenarios, supporting team sizing, monitoring of emerging demands, and time consumption analysis. The results indicate greater analytical efficiency and decision-making support with the use of the tool.

KEYWORDS

Topic Modeling, Helpdesk, Large Language Models

1 INTRODUÇÃO

Sistemas de *helpdesk* desempenham papel fundamental na prestação de serviços de suporte técnico em empresas de tecnologia, registrando milhares de ocorrências por meio de textos descritivos escritos por usuários. Apesar de conterem informações valiosas, essas descrições são difíceis de serem analisadas automaticamente [8].

Neste trabalho, apresenta-se uma ferramenta que integra estratégias de Processamento de Linguagem Natural (PLN), em particular a Modelagem de Tópicos (MT) [5, 13] e os Grandes Modelos de Linguagem (LLMs) [3], para organizar semanticamente informações de *helpdesk* e apoiar o processo de tomada de decisão. A ferramenta é composta de três módulos: (i) um que coleta os dados, integra com sistemas de *helpdesk* e realiza um pré-processamento que padroniza as informações textuais por meio do uso de LLMs; (ii) um que recebe as informações padronizadas e aplica estratégias de MT para organizar os dados e extrair informações; e (iii) metáforas visuais que auxiliam na identificação de padrões.

Avaliamos a ferramenta por meio de um estudo de caso aplicado a dados de *helpdesk* de uma empresa de tecnologia da informação, contendo um total de 9.015 *tickets*. A aplicação possibilitou aprimorar o dimensionamento da equipe interna e externa, prever demandas emergentes e analisar o consumo de horas por tipo de atendimento. Além disso, a criação de metáforas visuais permitiu

que gestores identificassem tendências e tomassem decisões mais embasadas sobre a alocação de recursos.

2 TRABALHOS RELACIONADOS

A MT tem se destacado como uma abordagem não supervisionada eficaz para a organização e extração de informações em grandes volumes de dados textuais. Na indústria de software, MT tem sido aplicada para extrair características e problemas em avaliações de usuários, contribuindo na priorização de melhorias em aplicativos [4, 8]. Em recomendação, tópicos extraídos de avaliações auxiliam na personalização de ofertas com base em preferências implícitas [2]. Em saúde, a MT tem sido empregada para explorar prontuários e artigos médicos, revelando relações entre sintomas, diagnósticos e tratamentos [14]. Em redes sociais, a extração de tópicos [15] identifica tendências e eventos emergentes a partir de textos curtos, como *tweets*. Apesar dos avanços, há uma lacuna na aplicação de MT em ambientes corporativos de suporte técnico, especialmente em sistemas de *helpdesk*, onde a análise de textos desestruturados pode fornecer *insights* estratégicos ainda pouco explorados. Este trabalho contribui nessa direção, aplicando MT para apoiar a tomada de decisão em ambientes de atendimento.

3 FERRAMENTA: DESENHO E ESTRUTURA

Nessa seção descrevemos a nossa proposta de ferramenta para análise de bases textuais de *helpdesk* que permite identificar automaticamente temas recorrentes ou emergentes, possibilitando a extração de *insights* estratégicos para a melhoria contínua dos serviços prestados. O objetivo é uma ferramenta que auxilie no dimensionamento de equipe, no monitoramento de demandas, otimizando o consumo de horas em cada tarefa. Ela é composta de três módulos, conforme ilustrado na Figura 1, detalhados nas próximas seções.

3.1 Coleta e pré-processamento

A coleta dos dados pode ser feita por APIs - *Application Programming Interface* [10], quando disponíveis, ou por exportação de arquivos em *csv*. Nossa proposta prevê a utilização das seguintes informações: (i) um identificador único de chamado; (ii) texto descritivo da demanda escrita pelo usuário; (iii) classificação (quando houver) do atendimento (e.g. um atendimento sobre construção, pode indicar se é da área de elétrica ou hidráulica); (iv) tempo gasto no atendimento; e (iv) data de abertura do chamado, para análises temporais.

As demandas descritivas, na maioria dos sistemas *helpdesk*, são geradas a partir de e-mails compostos de uma estrutura que contém os seguintes itens: uma **saudação** (e.g. "Bom dia", "Boa tarde",

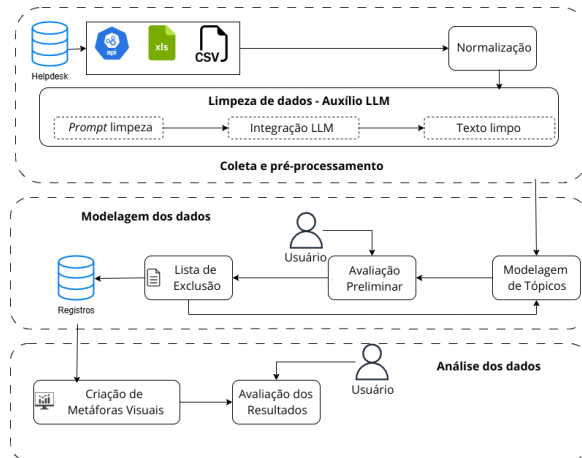


Figura 1: Funcionamento da ferramenta.

"Prezados", etc); o **texto demanda**; a **assinatura** e, eventualmente, **disclaimers**: (e.g. textos legais, indicando a confidencialidade, etc.).

Diante desta estrutura e da inexistência de um padrão definido para base de dados, para extrair as informações relevantes que interessam do ponto de vista de caracterizar e compreender do que se trata cada solicitação de atendimento, utilizamos um *prompt* para ser utilizado por LLM. Os grandes modelos de linguagem têm sido usados amplamente para realizar tarefas textuais dada a sua facilidade com a geração de textos [12]. No *prompt* (Tabela 1) é solicitado ao LLM que seja mantido apenas o texto da demanda, facilitando assim a extração do texto que realmente importa. Pelo elevado volume de dados e custo financeiro associado, optamos por um LLM de código aberto, o Llama3 [9]. Os dados tratados são armazenados em um arquivo csv enviado para o módulo seguinte.

Extract complete email signature from text. Return only text off output, equal of input.

All words in lowercase.

<input> Texto completo de exemplo

<output> Texto limpo

<input> {Evaluate Text}

<output> {Response from LLM}

Tabela 1: Prompt para Limpeza das assinaturas - LLAMA 3.

3.2 Modelagem dos dados

Para organizar semanticamente as demandas (textos) geradas por usuários de serviços de *helpdesk*, nossa ferramenta adota estratégias de Modelagem de Tópicos (MT), que visavam encontrar automaticamente tópicos semânticos a partir de documentos textuais. Encontramos na literatura uma vasta quantidade de propostas de MT, com destaque para duas estratégias consideradas estado da arte: **CluWords** [13] e **BERTopic** [5]. A técnica do CluWords substitui a representação TF-IDF [1] convencional por uma representação orientada à semântica [13], construída usando agrupamentos de palavras próximas em um espaço contextual de *word-embeddings* e, por fim, realiza a extração dos tópicos por meio de uma técnica de fatoração de matrizes (i.e. NMF). O BERTopic aproveita *embeddings* gerados pelo BERT [7], que permitem capturar nuances semânticas mais refinadas. Inicialmente, o modelo gera *embeddings* de sentenças que são reduzidas dimensionalmente, seguidas por um

agrupamento através de um algoritmo de agrupamento (i.e. HDBSCAN). A extração de palavras-chave dos tópicos é realizada por meio de um procedimento TF-IDF baseado em classes, melhorando significativamente a coerência dos tópicos extraídos. Em nossa ferramenta, adotamos as CluWords, por terem apresentados resultados mais consistentes em avaliações anteriores [11].

Após o processo de MT, nossa ferramenta permite uma avaliação prévia dos tópicos, que deve ser realizada por um especialista, para identificar palavras que eventualmente não têm relevância e que foram associadas aos tópicos. Além disso, nessa etapa prévia, esse especialista tem a possibilidade de criar identificações macro para cada um dos tópicos que, associadas às principais palavras dos tópicos, permitirá uma visão mais ampla dos problemas, por exemplo: tópicos identificados como "memória, hd, disco", podem ser classificados como "infraestrutura" e subclassificadas como "hardware", criando mais duas camadas de análise que permitem ao tomador de decisão identificar em um nível maior o problema e descer para outros níveis se desejar detalhar mais. Ao final desta etapa, os dados resultantes são armazenados para que possam ser consultados e utilizados por metáforas visuais que auxiliarão na tomada de decisão.

3.3 Interface Gráfica

Para permitir que análises sejam feitas a partir das informações extraídas na etapa anterior, propomos metáforas visuais que utilizam os tópicos encontrados e suas respectivas classificações utilizando uma ferramenta de *Business Intelligence* [6]. Neste caso o *PowerBI* da Microsoft foi escolhido pela empresa do estudo de caso. Os usuários podem realizar análises globais sobre os tópicos, até o detalhamento dos *tickets* associados a cada um deles, gerando confiabilidade e rastreabilidade aos dados. Todas as análises podem ser realizadas em janelas temporais distintas definidas pelos usuários, combinando com as informações estruturadas. Na Figura 2, ilustramos uma das metáforas da ferramenta, na qual o usuário consegue definir o período de análise (canto superior esquerdo). Os títulos dos tópicos associados a esse período são apresentados abaixo e a descrição do tópico e a nuvem de palavras associadas a ele são apresentados na direita, de acordo com o tópico selecionado (nas barras azuis).

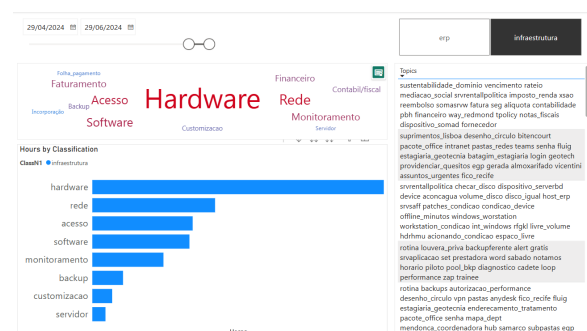


Figura 2: Ilustração da uma Metáfora Visual da Ferramenta

A ferramenta será licenciada sob a *Creative Commons BY*, permitindo seu uso livre desde que seja atribuída a autoria original, e está disponível em <https://github.com/paper-hd/paper>.

4 AVALIAÇÃO EXPERIMENTAL

4.1 Estudo de Caso

Para avaliar a aplicabilidade da ferramenta proposta, consideramos os dados de uma empresa de consultoria em TI que oferece serviços de suporte em infraestrutura tecnológica, desde manutenção de servidores e *firewalls* até apoio a usuários em aplicações da suíte Microsoft e suporte especializado em sistemas de gestão empresarial (ERP) utilizados por médias e grandes organizações. A equipe é composta por analistas especializados por módulos (financeiro, contábil, fiscal, entre outros) e utiliza um sistema de *helpdesk* que recebe demandas via e-mail ou portal, classificando-as entre Infraestrutura e ERP. Esta é a única classificação na qual um atendimento é atribuído atualmente. O período da base de dados considerado nessa avaliação está compreendido de janeiro de 2023 a julho de 2024 e foi dividido em trimestres para que o volume de *tickets* fosse satisfatório para as análises. No total, a base de análise contou com 9.015 *tickets*, conforme descritos na Tabela 2, separados por trimestre. Além disso, para nossa análise, foram tratados os campos indicados e representados na tabela 3. Durante a avaliação, a empresa contava com 10 analistas e atendia mais de 1000 tickets mensais para cerca de 30 clientes ativos. Os dados foram exportados do *helpdesk* em arquivo csv.

Referência	Número de Registros	ERP	Infraestrutura
01T2023	1.424	628	796
02T2023	1.714	756	958
03T2023	1.659	738	921
04T2023	1.290	613	677
01T2024	1.261	718	543
02T2024	1.667	1.067	600
Total	9.015	4.520	4.495

Tabela 2: Base de Dados

Campo	Tipo	Observação
Class N1	nvarchar	Indica a classificação do <i>ticket</i> entre Infraestrutura e ERP
ID	int	ID do <i>ticket</i> na base do cliente
Texto	text	Texto do <i>ticket</i> com detalhamento da solicitação
CreateTime	dateTime	Indica a data de criação do <i>ticket</i>
TotalTimeTracked	int	Indica o tempo, em segundos, gastos para a solução do <i>ticket</i>

Tabela 3: Campos da base de dados

4.2 Resultados Práticos

Dimensionamento de equipe: Em janeiro de 2023, a empresa, com base em uma análise sem a ferramenta, resolveu contratar uma analista de infraestrutura de nível 2, cujas responsabilidades e conhecimentos envolvem demandas com servidores, *firewalls* e outras tecnologias que necessitam de uma maior senioridade. Esse profissional, cuja faixa salarial é de aproximadamente R\$3.500,00 foi contratado, pois na percepção da gestão da empresa problemas com servidores eram recorrentes e necessitavam deste tipo de profissional. Em julho de 2024 esse analista foi desligado da empresa por falta de demanda.

Revisitamos essa situação considerando que a ferramenta proposta fosse utilizada na análise e os resultados são apresentados na Figura 2. Podemos identificar que, desde janeiro de 2023, nossa ferramenta fornece indícios claros que a maioria das demandas da empresa não estaria ligadas a um analista de infraestrutura de nível 2. Os tópicos encontrados na MT para o agrupamento de infraestrutura foram claros, destacando itens como: Hardware, Pacote Office; Login; Senha; Teams; e Suprimentos. Estes tópicos, conforme a especialista da empresa, são atividades de um recurso menos qualificado (e também com custo menor), sendo que, analisando detalhadamente os chamados de suporte destes tópicos, a maioria se trata de lentidão de computador, instalação de *drivers* ou de programas e suporte em ferramentas do cotidiano da empresa (Tabela 4).

ID Ticket	Descrição
1076	prezados, solicito uma avaliação do meu computador, pois o teclado está com algumas teclas sem funcionamento , além de estar muito lento . att,
1093	bom dia! não consigo acessar o site do crea sp.
1094	oi pessoal, tudo certo? peço por gentileza instalacao do protheus para o usuario mateus pimenta.
1104	prezados, boa tarde. solicito a possivel troca do meu notebook, pois o mesmo tem estado lento e travado diariamente afetando minha produtividade.

Tabela 4: Exemplo *tickets* de infraestrutura

A empresa conta com analistas pouco qualificados quando o assunto é muito especializado nas áreas de contabilidade, fiscal e departamento pessoal. Para estes casos, analistas externos são contratados para auxiliar nas demandas. Existe uma grande dificuldade para a contratação de analistas especializados, pois o custo é alto, reforçando a importância da ferramenta para auxiliar no dimensionamento da equipe. Para ilustrar esse cenário, simulamos uma análise de dados relacionados ao quarto trimestre de 2023, conforme apresentado na Figura 3. Nesse semestre, conforme relato da empresa, houve a contratação de um especialista na área contábil. Todavia, avaliando os insumos fornecidos pela ferramenta, os tópicos relativos ao departamento pessoal apareceram mais vezes que os tópicos de contábil/fiscal. Essas análises reforçam a importância da ferramenta para empresas do setor de *helpdesk*.

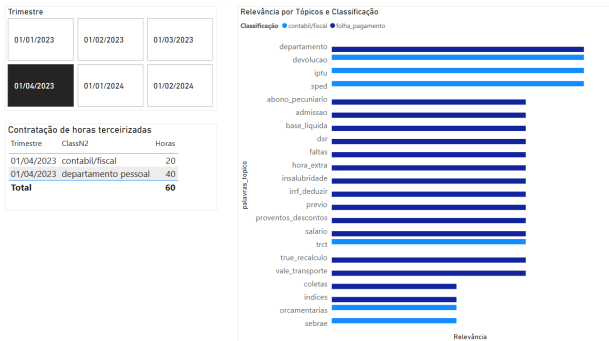


Figura 3: Tópicos ERP - Contábil/Fiscal e Dep. Pessoal - 4º Trimestre 2023

Consumo de horas de atendimento: A análise de consumo de horas é um grande desafio. Na concepção atual da empresa, existem apenas duas macro classificações prévias (Infraestrutura e ERP) e, conseqüentemente, essas análises são restritas a essas duas classes. Para ilustrar esse cenário, consideramos os dados relacionados ao tickets do segundo trimestre de 2024, com a ocorrência de 1067 *tickets* na categoria ERP, consumindo 557 horas, e 600 *tickets* na categoria infraestrutura, consumindo 586 horas. Quando essas informações foram reorganizadas de acordo com os tópicos automaticamente gerados por nossa ferramenta, conseguimos realizar uma análise mais específica de quais tipos de atendimento demandam mais tickets e horas, conforme apresentamos na Tabela 5

Pela análise de ERP, observamos que uma alta demanda de *tickets* pode não refletir na quantidade de horas gastas. A exemplo, o contábil/fiscal teve muitos *tickets* (247) e poucas horas de execução (55), sendo o mesmo refletido no financeiro (248 tickets para 24 horas) e relatórios (169 tickets para 86 horas). Por outro lado, focando nos dados da classificação de infraestrutura, claramente é possível concluir que a troca de dispositivos de hardware (120 tickets para 239 horas) e problemas de software (80 tickets para 133 horas) consomem mais tempo que as demais demandas que, por sua

vez, corroboram com as análises apresentadas que apontaram essas como sendo a de maior demanda para contratação de profissional.

ERP	Qtde	Horas	Infraestrutura	Qtde	Horas
contabil/fiscal	247	55	acesso	82	77
departamento pessoal	210	200	backup	50	25
faturamento	124	114	hardware	120	239
financeiro	248	24	monitoramento	82	55
incorporação	44	35	rede	104	101
obras	25	43	servidor	28	9
relatórios	169	86	software	133	80
Total:	1067	557	Total:	600	586

Tabela 5: Horas/Tickets por Macro Classificação e Tópicos

Previsão de Demandas Emergentes: A previsão de demandas é essencial para a correta alocação de profissionais, identificação de sobrecargas e redução do tempo de espera nas solicitações. Para auxiliar nessa tarefa, a ferramenta proposta permite organizar os dados em janelas temporais de 30 dias, associando os tópicos gerados às macroclassificações, além de possibilitar a execução diária de novas modelagens que se somam às anteriores, proporcionando uma visão de demandas emergentes. O objetivo dessa abordagem é prover agilidade ao setor operacional que não precisará analisar individualmente cada *ticket* para identificar tendências. Isso economiza tempo e permite uma leitura mais estratégica do cenário atual da empresa. Ao aplicar filtros nos últimos 30 dias, evidenciam-se os tópicos mais relevantes, proporcionando uma previsibilidade, facilitando a tomada de decisões baseada em evidências.

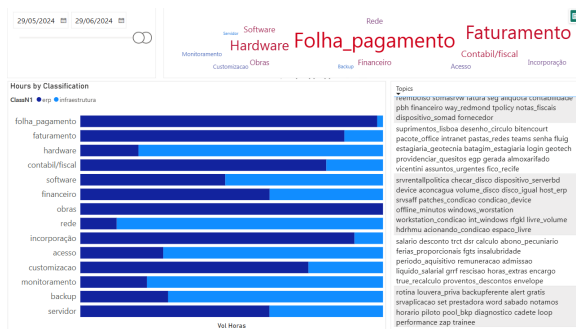


Figura 4: Tópicos últimos 30 dias



Figura 5: Detalhando Folha de pagamento nos últimos 30 dias

Para ilustrar essa funcionalidade, apresentamos na Figura 4 a análise provida pela ferramenta na qual avaliamos exclusivamente os tópicos mais relevantes relacionados ao período de 29/05/2024 à 29/06/2024, simulando o que um gerente da empresa poderia fazer antes do início do mês de julho para realizar uma prévia de demandas. Conforme podemos observar, fica evidente que demandas de

folha de pagamento são maiores que as demais. Para identificar quais os tópicos e quais palavras estão mais presentes, a ferramenta permite que seja clicado em folha de pagamento e os dados sejam filtrados, conforme ilustrado na Figura 5. Com este filtro, fica fácil de identificar as palavras-chave contidas nos *tickets*, sendo elas abono, admissão, entre outros, que identificam as demandas emergentes.

5 CONCLUSÕES E TRABALHOS FUTUROS

O artigo propôs e avaliou uma ferramenta baseada em MT e LLMs para extrair informações relevantes de grandes volumes de dados textuais oriundos de sistemas de *helpdesk*. A solução demonstrou eficácia na análise de dados não estruturados, contribuindo para uma gestão mais eficiente de equipes e recursos, ao fornecer *insights* sobre consumo de horas, previsão de demandas e identificação de padrões emergentes. Os resultados indicam, também, maior eficiência analítica e suporte à tomada de decisões com uso da ferramenta. Em trabalhos futuros, pretendemos incorporar técnicas mais avançadas de PLN, como o uso de LLMs para automatizar a extração e validação de tópicos. Também pretendemos integrar dados multimodais, visualizações interativas e adaptações da ferramenta a diferentes domínios, como atendimento ao cliente e análise de *feedback*, visando tornar o sistema mais robusto e estratégico, ampliando sua aplicabilidade em variados contextos organizacionais.

AGRADECIMENTOS

Trabalho apoiado por CNPq, INCT-TILD-IAR, AWS e FAPEMIG.

REFERÊNCIAS

- [1] Prafulla Bafna, Dhanya Pramod, and Anagha Vaidya. 2016. Document clustering: TF-IDF approach. In *2016 ICEEOT*. IEEE, 61–66.
- [2] Jesús Bobadilla, Fernando Ortega, Antonio Hernando, and Abraham Gutiérrez. 2013. Recommender systems survey. *Knowledge-based systems* 46 (2013).
- [3] Yupeng Chang, Xu Wang, Jindong Wang, Yuan Wu, Linyi Yang, Kaijie Zhu, Hao Chen, Xiaoyuan Yi, Cunxiang Wang, Yidong Wang, et al. 2024. A survey on evaluation of large language models. *ACM TIST* 15, 3 (2024).
- [4] Vinicius HS Durelli, Rafael S Durelli, Andre T Endo, Elder Cirilo, Washington Luiz, and Leonardo Rocha. 2018. Please please me: does the presence of test cases influence mobile app users' satisfaction?. In *SBES*, 132–141.
- [5] Maarten Grootendorst. 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794* (2022).
- [6] Richard T Herschel and Nory E Jones. 2005. Knowledge management and business intelligence: the importance of integration. *Jn. of knowledge management* (2005).
- [7] Jacob Devlin Ming-Wei Chang Kenton and Lee Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of naacl-HLT*, Vol. 1. Minneapolis, Minnesota.
- [8] Washington Luiz, Felipe Viegas, Rafael Alencar, Fernando Mourão, Thiago Salles, Dárlinton Carvalho, Marcos Andre Gonçalves, and Leonardo Rocha. 2018. A feature-oriented sentiment rating for mobile app reviews. In *WWW*.
- [9] Meta. 2024. *Introducing Llama 3.1: Our most capable models to date*. <https://ai.meta.com/blog/meta-llama-3-1/>
- [10] Joshua Ofoeda, Richard Boateng, and John Effah. 2019. Application programming interface (API) research: A review of the past to inform the future. *IJEIS* (2019).
- [11] Antônio Pereira, Felipe Viegas, Marcos André Gonçalves, and Leonardo Rocha. 2023. Evaluating the Limits of the Current Evaluation Metrics for Topic Modeling. In *Proc. the 29th WebMedia* 2023, 119–127.
- [12] Prabu Ravichandran, Jeshwanth Reddy Machireddy, and Sareen Kumar Rachakatta. 2024. Generative AI in Business Analytics: Creating Predictive Models from Unstructured Data. *Hong Kong Journal of AI and Medicine* 4, 1 (2024), 146–169.
- [13] Felipe Viegas, Sérgio Canuto, Christian Gomes, Washington Luiz, Thierson Rosa, Sabir Ribas, Leonardo Rocha, and Marcos Gonçalves. 2019. CluWords: exploiting semantic word clustering representation for enhanced topic modeling. In *WSDM*.
- [14] Hanna Wallach, David Mimno, and Andrew McCallum. 2009. Rethinking LDA: Why priors matter. *Advances in neural information processing systems* 22 (2009).
- [15] Xiaohui Yan, Jiafeng Guo, Yanyan Lan, and Xueqi Cheng. 2013. A bitern topic model for short texts. In *WWW*. 1445–1456.