Proceedings of the Nineteenth International AAAI Conference on Web and Social Media (ICWSM 2025)

A Growing Sense of Alienation: Spirals of Silence and Suppression of Structural Circumstances of Suicide in News

Jasmine C. Foriest[1*], Mini Jain[1*], Benjamin D. Horne[2], Munmun De Choudhury[1]

[1]Georgia Institute of Technology, USA
[2]University of Tennessee Knoxville, USA
jay@gatech.edu, mjain335@gatech.edu, bhorne6@utk.edu, munmund@gatech.edu

Abstract

Suicide is a leading cause of death in the United States. Global safe reporting guidelines for news reports of suicide intend to mitigate associations of increased suicide incidence and stigma. However, recent research suggests more latent patterns in news beyond the guidelines could still contribute to suicide outcomes such as inhibited help-seeking and isolation. Using the Theory of Spiral of Silence to center isolation, we take a mixed-methods approach to analyze 22,021 articles (2020-2024) and use a zero-shot learning large language model (LLM) classifier to detect suppression of four structural circumstances of suicide: financial/job, legal, school, and access to physical/mental healthcare. We find that circumstance disclosure by news publishers diverges by political leaning, financial ($p = 0.016$), legal ($p < 0.001$), and school ($p < 0.001$); and by regionality, legal ($p < 0.001$) and health ($p < 0.001$). We qualify mechanisms of suppression using topic modeling and content sharing networks (CSNs). The spiral of silence lens highlights that left leaning publishers are more likely to disclose systemically or socially collective circumstances. In contrast, right leaning outlets suppress those and instead disclose instances that blame individuals for their experiences. Our work highlights how news reporting can downplay structural factors contributing to suicide.

Content Warning: This paper discusses suicide deaths reported in news articles and may be sensitive to readers.

Introduction

Suicide is a pressing public health issue. Resulting in 49,476 deaths, suicide was the second leading cause of death in 2022 for ages 10-14 and 20-34 (NVSS 2024). Even more substantial is the estimated number of people experiencing suicide related behaviors in the same year. In 2022, 13.2 million Americans experienced suicidal ideation, 3.8 million planned to attempt suicide, and 1.6 million persons did attempt suicide (SAMHSA 2021). The rates of both suicide deaths and suicide related behavior have risen in recent years (NVSS 2024), and require advancements in research and prevention strategies – designated as an urgent threat and research priority for the Centers for Disease Control and Prevention's National Injury Center (Centers for Disease Control and Prevention 2023b).

Unsafe reporting in news media, as one of many factors, has been identified in research as a considerable influence on spikes in suicide (Centers for Disease Control and Prevention 2023a). Specifically, research shows that sensationalized reporting, like that of celebrity suicides, is positively correlated with increased suicide incidence (Phillips 1974; Sinyor et al. 2021). Attributes of news reports that contribute to this include descriptions of suicide epidemics or presentation of suicide myths (Niederkrotenthaler et al. 2010). While these findings have informed prevention strategies, such as the development of safe reporting guidelines, their efficacy (Pirkis et al. 2006), and journalistic adherence to them (McTernan et al. 2018; Roth et al. 2020; Sumner, Burke, and Kooti 2020), more work is needed to understand less easily enumerated attributes of news reports. Moreover, recent work has shown that stigmatizing and glorifying language persists in news reports of suicide deaths (Foriest et al. 2024) suggesting latent patterns like these in news contribute to unsafe reporting and preclude suicide prevention.

Essential to suicide prevention are help-seeking behaviors – which are connected to an individual's sense of isolation. The social environment, a lack of social support, and social isolation contribute to suicide risk and are considered in suicide etiology and theory (Durkheim 2005; Joiner 2005). While the definition is broad within suicidology, a working understanding provides that:

> "social isolation is a wider concept than living alone. It includes: the social and cultural isolation of the immigrant; the solitude of old age arising from a lack of contemporaries to share values and outlook; the unemployed's sense of social rejection; the ostracism resulting from infringement of a social taboo by divorce or criminal act, or any similar activity that might diminish relatedness to the community. A high suicide rate is found in all these categories; only the concept of social isolation embraces and accounts for such a diversity of phenomena (Sainsbury 1955)."

As Sainsbury illustrates, social isolation manifests in many forms and has many vehicles in which individuals experience othering – circumstances. Circumstances of suicide, negative life events that precipitate suicide-related behavior (Chen and Roberts 2021), are leveraged across suicide research to inform prevention. Structural circumstances, are those relating to interactions with systems and policies, such as the economic, legal, political, and healthcare systems (Castillo, Hansen, and Rocha 2018). And while these circumstances, along with social isolation, are included in existing frameworks as factors of suicide (NCIPC

[*]These authors contributed equally.

2022), some scholars contest that the factors are not thoroughly included in suicide research and that this exclusion enables oppressive systems to remain impervious to prevention strategies (Fitzpatrick 2018). Thus, this work prioritizes structural factors and isolation as components of this study in order to contribute to bridging the current gaps. In doing so, we answer two research questions:

RQ1: a) Can we detect disclosure of structural factors in news reports of suicide? b) How are the detected disclosures of structural factors in news reports of suicide distributed across news publishers in terms of their (a) political leaning and (b) regionality?

RQ2: How do we understand patterns of disclosure of structural factors within these groupings as suppression?

Towards these research questions, employing machine learning techniques on a large-scale and broad news dataset, we first identify articles about suicide involving a structural circumstance. We build a large language model (LLM)-based classifier for discerning the presence of 4 structural factors in news reports of suicide, which are Financial/Job Problem, Legal Problem, School or Academic Problem, and Lack of Access to Health/Mental Health Care. We then compare the distribution of structural factors of suicide across political and regional axes of comparison using statistical analysis. Then, we dive deeper into each structural factor and use topic modeling to study how different themes within each detected structural factor are talked about amongst different political and regional news groups. Then, we look at how these differences are amplified through patterns of copycat behaviors amongst news communities.

We find that there is a significant difference between the distribution of structural factors across political and regional axes of comparison. Our results from topic modeling indicate differences in patterns of reporting of each structural factor, which show that left-leaning publishers have higher probabilities of disclosing structural circumstances systemically or from a social perspective whereas right-leaning ones talk about it at the individual level, suppressing the systemic issue. Further, these differences are amplified through patterns of news copying/syndication, where we find the most copying happening when the structural circumstance is a legal problem. Through these results, our work sheds light on the muting of structural factors of suicide in news reporting.

Contributions. Taken together, our paper makes two contributions – one theoretical, one methodological:

- Our *theoretical contribution* comes from drawing on the Spiral of Silence theory (Noelle-Neumann 1974), described in the next section, as a theoretical framework. We contribute a novel operationalization of this communications theory in the social sciences in the study of mass media's influence on public opinion related to the public health outcome of suicide.
- Our second contribution is *methodological*. While many studies across the fields of computational social science, communications, and public health have contributed to the topic of suicide, this work to the best of our knowledge is the first to leverage LLMs or an unsupervised approach to discern the relationship between news media and circumstances of suicide. This contribution is intended for multidisciplinary audiences of public health, suicidology, communications, and computational social science. We respond to recent calls for increased use of computational methods to support suicide prevention practitioners in automating the detection of unsafe reporting of suicide. Our demonstration of the efficacy of cutting-edge machine learning (LLMs), paired with classic and reliable methods such as topic modeling, in discerning suppression of structural factors of suicide can be readily leveraged by teams seeking to produce explainable tools and resources to lay audiences like journalists in executing safe reporting practices.

Related Work

Media Studies in Suicide Prevention

There has been ample amount of work in predicting suicide behaviours online, such as (Gollapalli, Zagatti, and Ng 2021) which uses topic modeling on the user's social media posts to predict the risk of suicide. A lot of recent work has focused on analyzing and detecting suicide ideation through natural language processing methods like RNNs (Coppersmith et al. 2018) and linguistic tools like SAGE (Sawhney et al. 2021) applied on the individual's social media footprint, with some work leveraging on the user context (Flek 2020). More recently, LLM based prompting methods have also been used for detecting evidence of suicide ideation in social media posts, by Singh et al. (2024) as part of the CLPsych 2024 Task focused on suicide ideation detection (Chim et al. 2024). There has also been some research into studying how social support found in online communities influences suicide ideation risk over time amongst individuals (Choudhury and Kıcıman 2017). Other work has explored the impact of news on suicide ideation, finding that reporting of celebrity suicides in news media led to an increase in suicide ideation and intent expressed on social media (Kumar et al. 2015a) as well as number of deaths by suicide (Niederkrotenthaler et al. 2020).

Though there is a breadth of work across communication, social computing, and computational social science (CSS) literature towards suicide prevention, this work is first of its kind to apply it to study suicide or suppression in news. Moreover, this work is, to the best of our knowledge, the first to leverage LLMs to discern the relationship between news media and circumstances of suicide.

Structural Factors of Suicide

Given the public health burden of suicide deaths in the U.S. (NVSS 2024) and incidence of suicide related behavior (SAMHSA 2021), identifying and addressing events or factors that precede suicide are essential to prevention. Public health research has developed frameworks to study and prevent suicide. The CDC organizes one framework as risk and protective factors of suicide (NCIPC 2022). Categories of factors are organized by level: societal, community, relationship, and individual (NCIPC 2022). Application of this framework has empirically associated risk factors with suicide death and ideation using network analysis (Holman and Williams 2022), protective factors like social support with decreased lifetime likelihood of suicide attempt, and intersections with other types of violence – shared risk and protective factors (Wilkins et al. 2018).

	#Articles	#Publishers	Avg. #Words	Left	Right	National	Regional
Articles after keyword filtering	22021	304	1119.88	9058	4184	14498	7066
Articles after LLM filtering	13629	279	5088.98	5651	2471	6977	4426
Articles relevant to structural circumstances	9256	263	5376.16	3806	1602	4668	3041

Table 1: Summary statistics for extended NELA datasets used in this study

However, some scholars argue that existing suicide prevention strategies, informed by frameworks like risk and protective factors, attribute responsibility to the individual, obscure the impact of the socioeconomic, the political, the structural, and deflect away from community responsibility (Fitzpatrick 2018; Chandler, Cover, and Fitzpatrick 2022; Dawson and Silva 2009). Fitzpatrick argues that this underwhelming amount of attention to structural contributors to suicide is unethical – further disadvantaging oppressed groups and freeing institutional agents from their role in adequately resourcing prevention (Fitzpatrick 2018). The research and practice issue of individual responsibility and minimizing structural context is further illustrated in the attribution of factors such as criminal/legal and financial/job being attributed at the individual level, discrimination and healthcare access at the relationship level (NCIPC 2022). None of those essential factors are imposed at the societal level – potentially leaving applications of the framework to miss the role of structural oppression as a backdrop of all factors. In response to this consideration, we center structural factors in our study, specifically, we inquire about meaningful differences in disclosure of these factors by news.

Theoretical Framework: Spiral of Silence

In order to investigate the latent linguistic patterns of news propagated dissuasion of help-seeking behavior that are essential to suicide prevention strategies, we leverage a theory at the intersection of news media and isolation. The Spiral of Silence is a theory of public opinion formulated by German communications researcher Elisabeth Noelle-Neumann (Noelle-Neumann 1974). Noelle-Neumann offers that the sphere of public opinion, at least the most dominant and prevalent of positions, come to be through a spiraling process in which individuals seek to avoid any degree of social or intellectual isolation by withholding their dissenting opinions (Noelle-Neumann 1974). With no contest, the dominant opinion grows louder while opposing perspectives shrink quietly away – the individuals holding them ironically fragmented within the public anyway.

We operationalize the of Spiral of Silence (Noelle-Neumann 1974) as a lens for characterizing patterns of suppression of disclosure regarding structural factors of suicide in news media. It characterizes established phenomena involving in isolating of perspectives or experiences through news media – a demonstrated medium for suicide prevention. We operationalize the five main hypotheses of Spiral of Silence (Noelle-Neumann 1974) into three core mechanisms of suppression. The first two of these hypotheses orient us to the process of individual experiences of structural factors, differing from the public opinion in the news they consume, constituting **isolation**. The next two hypotheses describe the mechanism of **divergence**: where distributions of disclosed factors across social environments are different. The final hypothesis concludes that in the case of isolation and divergence. an individual perspective has succumbed to **suppression** – decreased likelihood of expression and in this case suicide prevention, or help-seeking.

Data

We used three news datasets, a manually annotated sample of `All The News`, `NELA-GT`, and `NELA-Local`. `All The News` (Snapcrack 2021) dataset is publicly available, comprising $143,000$ English news articles from 27 American publishers between 2016-2020. The annotated sample was retrieved using suicide-related keywords (Choi et al. 2020) curated by expert public health researchers and practitioners in the field of injury prevention and suicide. These exact terms were also used in previous research of suicide circumstances of news (Foriest et al. 2024), informing the methodology for our study. The manually coded `All The News` consists of 200 articles reporting suicide deaths. Each article was manually coded by two human coders with substantial agreement for suicide circumstances using a validated annotation framework (Chancellor et al. 2021). For this study, we are interested in structural determinants of suicide and isolated articles annotated with the following factors: 'Financial/Job Problem', 'Legal Problem', 'School or academic-related Problem,' and 'Lack of Access to Health or Mental Health Care'. This yielded 13, 39, 5, and 7 articles for the Financial, Legal, School and Health categories respectively – a total of 64 articles. We used these human annotated articles to inform prompt engineering and as a ground truth for validating our models.

`NELA-GT` and `NELA-Local` are curated datasets of news articles published by U.S. national, local, and fringe media outlets labeled with metadata regarding each article's publisher, political leaning, regionality, and veracity. (Nørregaard, Horne, and Adalı 2019; Horne et al. 2022). Comparable to this study, both datasets have been used to study news, policy frames, and violence against women (Mittal et al. 2024).

We sample news data from extended versions of the `NELA-GT` and `NELA-Local` datasets. Specifically, our sample consists of news articles between January 2020 and February 2024 from 304 US publications and related metadata about the publisher including its political leaning and regionality. We filtered the dataset using the same set of suicide-related key terms (Choi et al. 2020) from (Foriest et al. 2024) to get all the articles talking about suicide, and the filtered result consisted of $22,021$ articles as described in Table 1, alongside other descriptive statistics given in the first row of the table.

Zero shot prompting						Few shot prompting					
Financial Problem											
Model	Acc	Prec	Rec	F1	AUC	Model	Acc	Prec	Rec	F1	AUC
GPT4	0.91	0.20	0.30	0.24	0.73	GPT4	0.93	0.25	0.33	0.29	0.63
GPT3.5	0.93	0.08	0.30	0.12	0.64	GPT3.5	0.87	0.06	0.08	0.07	0.50
Mixtral	0.74	0.10	0.33	0.15	0.55	Mixtral	0.85	0.13	0.23	0.17	0.56
Mistral	0.82	0.14	0.09	0.07	0.51	Mistral	0.45	0.07	0.32	0.12	0.47
LLama	0.67	0.13	0.66	0.22	0.62						
Legal Problem											
GPT4	0.93	0.80	0.89	0.84	0.92	GPT4	0.81	0.55	0.67	0.60	0.76
GPT3.5	0.86	0.67	0.67	0.67	0.79	GPT3.5	0.79	0.50	0.56	0.53	0.70
Mixtral	0.69	0.35	0.49	0.43	0.64	Mixtral	0.77	0.32	0.15	0.21	0.54
Mistral	0.76	0.22	0.34	0.28	0.55	Mistral	0.74	0.33	0.22	0.26	0.55
LLama	0.74	0.42	0.82	0.55	0.77						
School Problem											
GPT4	0.98	0.67	1.0	0.80	0.98	GPT4	0.88	0.29	0.44	0.32	0.83
GPT3.5	0.98	0.67	1.0	0.80	0.98	GPT3.5	0.91	0.25	0.50	0.33	0.71
Mixtral	0.90	0.17	0.80	0.29	0.85	Mixtral	0.84	0.47	0.70	0.55	0.77
Mistral	0.96	0.33	0.80	0.47	0.88	Mistral	0.69	0.18	0.89	0.30	0.77
LLama	0.83	0.14	0.50	0.22	0.71						
Health Problem											
GPT4	0.79	0.18	1.0	0.31	0.89	GPT4	0.65	0.12	1.0	0.21	0.82
GPT3.5	0.79	0.18	1.0	0.31	0.89	GPT3.5	0.60	0.11	1.0	0.19	0.79
Mixtral	0.65	0.07	0.71	0.13	0.68	Mixtral	0.86	0.08	0.29	0.13	0.58
Mistral	0.79	0.13	0.86	0.22	0.82	Mistral	0.60	0.10	1.0	0.10	0.47
LLama	0.51	0.08	1.0	0.16	0.74						

Table 2: Zero and few shot prompting results across models and structural factors.

RQ 1: Disclosure of Structural Factors in News Reports of Suicide

In this section, we describe the approach for developing a classification model to detect disclosures of structural factors of suicide in news – part (a) of RQ1.

Classification Methodology

Our initial task requires discerning when articles are (a) relevant to suicide or suicide-related behavior and (b) attributing at least one of the four structural circumstances, per RQ1(a). For that purpose, we leverage the manually annotated sample of 200 articles from `All The News`. This dataset served as a valuable ground-truth for engineering our prompt for a classifier based on LLMs, and validating the performance of our classifier.

We formulated the problem of classifying each of the articles as belonging to none (zero) or more structural factors in the following way. Given a news article, it can be classified on 2 levels : (a) *Does it talk about a death by suicide?* (b) *Does it align with any of the 4 structural factors as a cause of the discussed suicide? If yes, then why?* We used LLMs with zero-shot and few-shot learning to build classifiers for this task, and also generate *rationales* for why a structural factor is detected in a news report of suicide. We choose this approach of using LLMs for classification, given the limited number of labeled samples in the annotated `All the News` dataset and the ability of LLMs to perform well on natural language understanding tasks without fine-tuning.

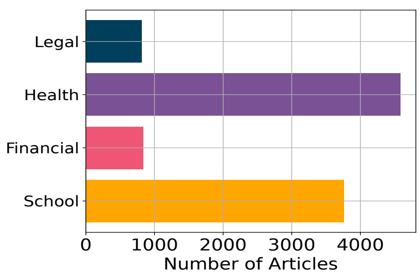
Zero-shot learning. Zero-shot learning (ZSL) (Kojima et al. 2022) (Liu et al. 2023) is a technique in which we use pre-trained models trained on large amounts of textual data from various domains, to answer natural language queries. This does not require training of the model or annotated samples.

Few-shot learning. Few-shot learning (FSL) (Brown et al. 2020) uses pre-trained large language models by providing it n demonstrative examples (n-shot learning) of the type of queries that the model is expected to answer, to "train" the model to answer in the expected manner. This allows us to incorporate additional knowledge into the model answers without explicitly doing any iterations of training.

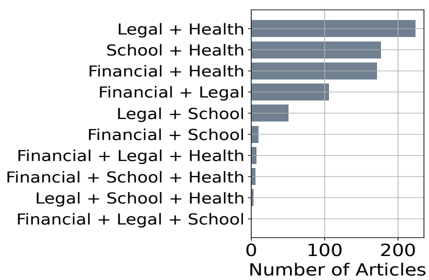
We applied zero-shot and few-shot learning using the following as base LLMs for comparative purposes. We chose these models due to their widespread usage and performance in LLM benchmarks.

- *LLaMa2*: LLaMa is a family of LLMs introduced in (Touvron et al. 2023). For our experiment, we use the 70B variant of LLaMa v2.
- *Mistral*: Mistral is a state-of-the-art open-source LLM by (Jiang et al. 2023). We use the Mistral-7B-v0.2 variant, which contains about 7.24 billion parameters.
- *Mixtral*: Mixtral is a pretrained generative Sparse Mixture of Experts by (Jiang et al. 2024). We use the Mixtral-7x8B-v0.1 variant, which contains about 46.7 billion parameters.
- *Generative Pre-trained Transformer (GPT)*: Generative Pre-trained Transformer (GPT) is a family of large language models released by OpenAI (Brown et al. 2020). For our experiment, we utilize the gpt-3.5-turbo-0613 with 175 billion parameters and the gpt-4-0613 model with 1.76 trillion parameters.

For developing the appropriate prompt, we split the `All The News` Dataset into a training and testing subset with an 80:20 split and passed the training dataset through multiple iterations of prompting on the `All the News` dataset. We found that the model performed the best when the language in the prompt was direct and consistent. Using a more literal expression for describing suicide, i.e. 'killed themselves' yielded better results. Explicitly adding all conditions that could be indicative of a structural circumstance in an 'or' format also helped improve the performance. We also generated rationales for why the model classifies a certain article into a particular category. This helped modify the prompt according to how the model was interpreting the text. We also looked at the rationales generated by the LLM for each classification to inform our prompting strategy and modify the language. The final prompt for ZSL can be found in Table A1 in the Appendix, which was modified for FSL by adding additional text with hand-curated examples of news articles; see Table A2.



(a) Single-Label



(b) Multi-Label

Figure 2: Results of ZSL using GPT-4

Results of Detecting Structural Factors

In response to RQ1, we find that our approaches – prompt-based classification – can successfully detect structural factors in news reports of suicide. Specifically, as shown in Table 2, using 20% of the ground truth of the human-annotated dataset as the unseen testing set, zero-shot learning models achieve a high AUC socre of 0.73, 0.92, 0.98 and 0.89 for Financial/Job Problem, Legal Problem, School or Academic Problem and Lack of Access to Health/Mental Health Care respectively.

Compared to the other models considered (e.g., FSL), we found GPT-4 consistently matches or outperforms GPT-3.5 on the human-annotated dataset as indicated by AUC. (Table 2). Hence we used this model to detect the presence of structural factors in the NELA dataset articles for our analyses using the same prompt. For an extended discussion of the relative comparison of the models, readers are directed to the Appendix.

From the results of the ZSL classification (Figure 2a), we obtain $13,629$ articles as actually talking about a death by suicide, out of which 9256 articles (67.91%) are about structural circumstances. The distribution of the structural factors among those articles was 810 (8.75%) classified as a Financial/Job Problem, 4579 (49.47%) classified as a Legal Problem, 834 (9.04%) as a School or Academic Problem and 3756 (40.58%) as Lack of Access to Health/Mental Health Care. Additionally, there are 707 articles that have been classified as belonging to more than one factor. Amongst all, Lack of Access to Health Care is the most commonly co-occurring one, occurring 556 times in multi-label classification (Figure 2b), which aligns with public health knowledge that these factors often occur together. This suggests that lack of access to healthcare is often exacerbated or accompanied by other structural factors.

Error Analysis of ZSL Classifiers

The validation step of the analysis, using the annotated dataset to get the results of ZSL with GPT-4, revealed some issues of misclassification that we reflect on. Discerning challenges of misclassification is essential to inform leveraging automated, unsupervised approaches for support tools to address unsafe reporting.

There were a total of 76 misclassifications. False-positives were highest in the 'Health Problem' category with 50 while the second highest were from the 'Legal Problem', with 26. Below, we provide the rationales generated by the LLM to illustrate the context of error.

Presence of a health problem is the absence of healthcare. The primary reason for misclassifications in the 'Health Problem' category, short for 'Lack of Access to Healthcare,' is that the LLM appears to occasionally infer that the presence of a physical health or mental health problem means an absence of treatment or care:

> "It is mentioned that she had emotional turmoil and struggled with her sexuality throughout her life. She also had "deep depressions" and her language seemed to "presage the struggles to come," possibly manic episodes "suffused with velocity and energy until she plunged into depression and despair." This indicates that she was dealing with mental health issues, which led to her suicide."

This example contrasts the instances in which the LLM appropriately identifies lack of access to care, denial of care, or ineffective care as a structural circumstance:

> "The article mentions a person who committed suicide due to the severe and continuous medical complications she faced after a mild COVID infection. It also talks about the struggles of long COVID patients who are unable to access proper healthcare or treatment, which aligns with the 'Health Problem' category."

Distribution of Disclosure Across Publishers

The second part of RQ1 permits us to investigate any differences between opinion groups or social environments. To understand how reporting of structural factors as a cause of suicide differs across news media, we analyzed the distribution of each of four classified structural factors in the news articles across two axes of comparison: (a) political leaning and (b) regionality. We then discerned the statistical difference in the attribution of each factor between corresponding comparison pairs. We used Welch t-tests for comparisons of

	#Left	#Right	p-value	t-statistic
Financial	300	166	**0.016**	-2.407
Legal	1983	759	**0.000**	3.890
School	320	171	**0.035**	-2.110
Health	1471	632	0.667	0.431
	#National	#Local	p-value	t-statistic
Financial	392	274	0.209	-1.257
Legal	2390	1317	**0.000**	5.046
School	394	283	0.105	-1.624
Health	1824	1453	**0.000**	-7.594

Table 3: Comparison across political leaning and regionality

left vs. right and national vs. local. This test allows accounting for the imbalance of articles across the axes, as detailed in Table 3.

This analysis revealed differences in coverage between both political leaning and regionality. We found that there is a significant difference ($p \leq 0.05$) in the occurrence of Financial (0.016), Legal (0.00010) and School (0.035) problems in left and right-leaning articles, and a significant difference in the occurrence of Legal (0.00000046) and Health ($p = 3.4 \times 10^{-14}$) problems in national and local articles. Left-leaning publishers have greater disclosure of financial/job and school problems preceding suicide deaths. Whereas the right talks more about problems related to legal issues. This is different in the comparison between distributions of structural factor disclosures between national and local publishers. Legal problems are talked about in greater frequency in national news, and health problems are talked more in local news.

RQ 2: Patterns of Suppression in Disclosure of Structural Factors Within These Groupings

After discerning the difference in the disclosure of certain structural factors by publishers with different political leanings and regionality, we contextualize the degree to which divergence in disclosure constitutes suppression of experiences to answer RQ2. We did this by first focusing on the factors for which there is a statistically significant difference in disclosure: for political leaning that is financial/job, legal, and school and for regionality that is legal and health. We analyzed suppression patterns in two ways that allow us to discern context and mechanism. The first was topic modeling and the second was the use of news communities.

Methods

Topic Modeling of Detection Rationales. Our first approach utilized the rationales we sought from the zero-shot learning classifiers during the detection task in RQ1. We assumed that by understanding *how* the models identified certain structural factors to be present or absent in a news article, we would be able to gain insight into suppression of specific factors in specific articles.

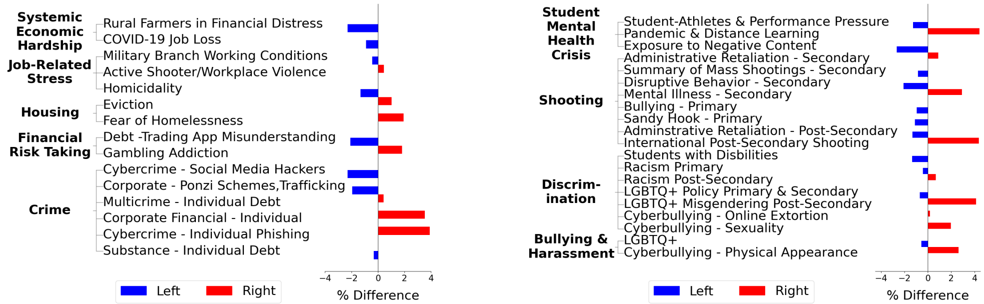
For this, we chose BERTopic (Grootendorst 2022) as our topic modeling approach, a transformer-based language model that extracts highly coherent topics on the rationales produced by the LLM. The input to these topic models were the rationales provided by the zero-shot learning classifiers for each of the structural factors.

We created a BERTopic model for each structural factor, using the Sentence-BERT "mpnet-base-v2" model (Song et al. 2020) to generate embeddings for the LLM rationales. We experimented with generating different number of topics for various instances of the topic models, ranging from 10 to 50, aiming to maximize the Coherence Score per factor – just as performed in similar work (Oghaz et al. 2020) to optimize topics within tweets. Then, with the optimal topic model, we obtain the distribution of topics by calculating the posterior probabilities for every representative rationale. We subsequently employed qualitative labeling informed by our domain expertise to identify broader, semantically interpretable themes from the resulting topics. With the probabilities hidden, we reviewed the representative documents and tokens for each topic to discern a theme that situated each topic within the structural factor and a broader social context. Using the posterior probabilities, we averaged per-topic probabilities by political leaning and regionality and then computed the % difference for each topic across our access of comparison. A negative probability for a topic in our analyses indicates that right-leaning articles are more closely related to that topic than left-leaning articles and vice-versa for a positive probability. The same extends to regionality as well, with a negative probability indicating more alignment with local publications and a positive probability meaning the topic is more similar to the national publication articles.

We performed semi-open coding during qualitative analysis of representative tokens and documents generate topic labels and sub-labels, categories and subcategories, to better contextualize topics. This was done leveraging two human raters' domain expertise in public health, social determinants of health, and suicide. Probability differences between comparison groups for each topic were analyzed only after all topic labels were determined to prevent confirmation or anchor bias.

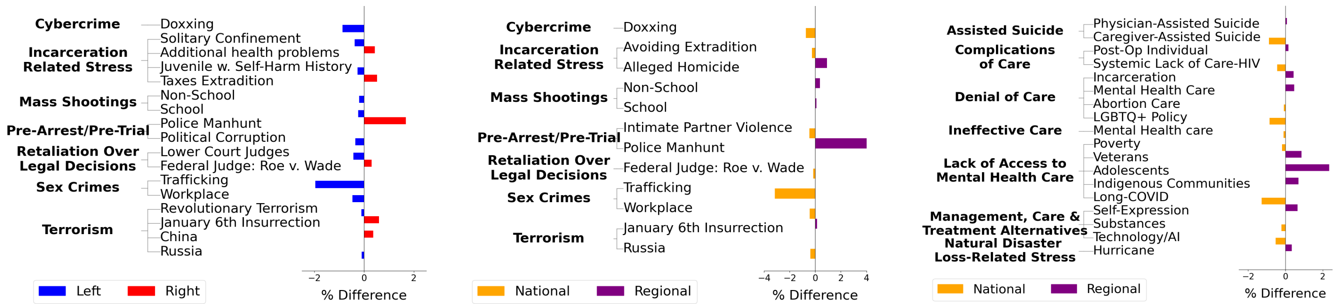
News Communities. The second approach involved discerning patterns of suppression in the mechanisms of news sharing or copying within latent publisher communities. Specifically, we constructed a Content Sharing Network (CSN) over all articles in the NELA datasets, whether than were related to suicide or not, using the algorithm developed in Horne, Nørregaard, and Adalı (2019). In short, we captured articles that are near-verbatim copies of each other by taking the cosine similarity greater than 0.85 of TFIDF vectors over 5-day windows of the dataset. Using these pairs of copied articles, we created n directed networks, where n is the number of unique sets of copied articles, by ordering articles by publication timestamps ($A \rightarrow B$ means B copied from A). Then we pruned these networks into n directed trees, keeping edges between articles with highest cosine similarity and breaking ties by article age (similar to Kruskal Algorithm for Minimum Spanning Trees (Kruskal 1956)). Finally, these article nodes were aggregated into outlet nodes to form the final network. For more details on this method, please see description in Horne, Nørregaard, and Adalı (2019). This general method has been used to analyze narrative amplification by fringe news media in the context of Syria Civil Defence (Starbird et al. 2018) and to examine differences in reporting across publisher communities in the context of Violence Against Women (Mittal et al. 2024).

Using this CSN, we investigated the presence and absence of copying suicide-related articles within communities to contextualize instances of amplification and suppression.



(a) Topics within Financial (Political)

(b) Topics within School (Political)



(c) Topics within Legal (Political)

(d) Topics within Legal (Regional)

(e) Topics within Health (Regional)

Figure 3: Subplots (a) Financial/Job Problem (b) School or Academic Problem and (c) Legal Problem show topic-wise percentage difference in probabilities by political leaning. Subplots (d) Legal Problem and (e) Lack of Access to Health/Mental Health Care show topic-wise percentage difference in probabilities by regionality. Confidence intervals are not provided for this subset of topics. Probabilities for all topics are located in appendix Tables A3-A6

Importantly, this analysis likely under-counts the spread of certain stories as it only captures near-verbatim copies of articles rather than paraphrases or other types of story coverage. However, this method provides concrete evidence of story amplification or lack thereof without capturing an excessive amount of noise. In Figure 4a, we show the network colored by communities as determined by modularity in Gephi (Bastian, Heymann, and Jacomy 2009). In this figure, we annotated the communities that most frequently copied suicide-related articles. The distribution of copied articles across these communities is shown in Figure 4b, and example outlets are shown in Table 4. Lastly, we show the number of copied suicide-related articles grouped by media type, structural category, and topics in Figure 5.

RQ2 Results

In this section we characterize the topic categories and subcategories with the most compelling illustrations of suppression by comparison group and then by divergent factor. We report the results of the topic models in Figure 3, but we only visualize the topics with the greatest differences in probabilities or disclosure patterns. The detailed results of the topic models can be found in Tables A3, A4, A5 and A6. We also describe the content copying dynamics within news communities and how they may contribute to suppression.

Political Leaning. Our topic modeling results allow understanding of suppression as reporting proclivities toward framing of individualism, individual culpability, or one-off occurrences for structural circumstances by right leaning publishers. That is contrasted by collectivist, population level, advocacy, and reformative perspectives of left leaning publishers. This holds for all factors. We elaborate below.

Financial/Job. The rationales of the zero-shot learning model for articles disclosing a financial or job problem preceding a suicide death or behavior yielded 19 topics. Our analysis of the topics produced 5 broad categories of financial or job problems in which individual topics were placed: Crime, Financial Risk Taking, Housing, and Job Related Stress. These topic categories, and the subcategories within them, are visible in Figure 3.

Pertaining to disclosures of financial or legal problems in suicide reports, some topics appear exclusively covered by either type of publisher. Left leaning publishers had higher probabilities of disclosing this factor in the context of Systemic Economic Hardship, a population and structural issue experienced by many in the cases of COVID-19 (0.90%) or rural farmers in the agriculture industry (2.29%). This contrasts sharply to right leaning publishers covering all Housing related disclosures that impact a single individual or appear to be the result of an individual's financial irresponsibility – such as homelessness (-1.93%), eviction (-1.04%), or divorce (-0.26%).

This pattern persists in other categories in which both the left and right disclose financial or job problems, but they do so differently – as exhibited by higher probabilities. In the case of Financial Risk Taking, the right had higher probabilities of disclosing an individual's gambling addiction (-1.82%) compared to left leaning disclosures of a company's failure to properly inform and support users of a financial risk taking act (2.08%). Similarly for financial crime, the right had high probabilities of reporting when the crime was perpetrated by an individual toward an entity or institution, or might otherwise be the result of a personal failing – 1 to 1. Examples of these include corporate financial crime perpetrated by an individual (-3.53%) and victims of cybercrime phishing scams (-3.90%). However the left discloses when the financial crime is perpetrated by 1 to many or many to many– such as ponzi schemes and trafficking (1.95%) or multiple social media hackers and victims (2.30%).

Legal. The rationales of the zero-shot learning model for articles disclosing a legal problem preceding a suicide death or behavior yielded 34 topics and 8 broad categories of legal problems in which individual topics were placed: Cybercrime, Evading Prosecution, Pre-Trial/Pre-Arrest, Incarceration Related Stress, Mass Shootings, Retaliation Over Legal Decisions, and Terrorism. We find instances of left leaning publishers being more likely to disclose circumstances where an institution imposing a legal outcome exacerbates an individual's degree of isolation, vulnerability, and *incarceration related stress*. This is exemplified in the case of solitary confinement (0.38%) or juveniles with a history of self-harm (0.26%) – framing as a call for institutional accountability and reform. Similar to financial and job problems, each category of publishers was more likely to disclose legal problems exclusively in some categories. For sex crimes, in this dataset, only left leaning publishers had probabilities of disclosing. And while not exclusive, right leaning publishers dominated the probability in the case of disclosures related to terrorism (-0.60%, -0.12%, -0.37%). Topical probability differences between left and right publishers are smaller overall compared to other factors.

School. Rationales in the detection models of school problem yielded 24 topics. Our analysis of the topics produced 5 broad categories of school problems in which individual topics were placed: Bullying & Harassment, Discrimination, Shooting, Social Media, and Student Mental Health Crisis. There were no broad topics categories that were dominated by left or right leaning publishers. Instead, we observe a more specific tendency related to school setting and subtopic. Left leaning publishers had higher probabilities of disclosing any school problems when they were related to individuals or incidents in primary or elementary school. This is observed in subtopics within the school shooting category, such as Sandy Hook Elementary (1.09%) or because of bullying (0.94%), and the discrimination category in the cases of LGBTQ+ policy (0.69%) or racism (0.42%) impacting primary and secondary school students. Within the bullying & harassment category, right leaning publishers had higher probabilities of talking about all contexts of cyberbulling: physical appearance (-2.65%), sexuality (-1.98%), and online extortion (-0.20%). Finally, both publisher types have greater probabilities of disclosing issues of student mental health but they oppose in the attribution of the crisis. This is illustrated by a higher probability to disclose that pandemic distance (-4.42%), social and learning, is to blame for students mental health decline. While left leaning publishers were more likely to disclose when social and institutional pressure to perform incited diminished student athlete mental health (1.25%).

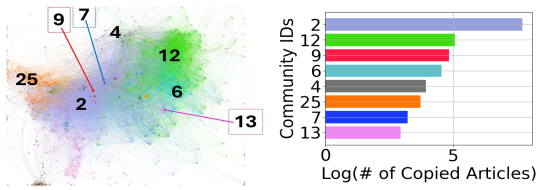
Regionality. The topic categories, subtopics, and their probabilities for national and local publishers are presented in Figure 3d and 3e. A rather consistent trend, comparable to differences between left and right publishers, is that local publishers have higher of probabilities disclosing in the cases of a local or state law where national publishers had higher probabilities of disclosing in the case of established social or political debate. Local publishers disclose when the parties affected can provide intimate personal context – national publishers are more likely to disclose in the case of celebrity.

Legal. National outlets cover topics with multi jurisdictional impact, such doxxing (0.73%) in the category of cybercrime, avoiding extradition (0.27%), terrorism in Russia (0.39%), retaliation over federal legal decision regarding Roe v. Wade (0.15%). In contrast, local outlets had higher probabilities of disclosing when there is a geographical or jurisdictional constraint on impact on the public. Such as a pre-arrest or pre-trial police manhunt (-1.68%), school mass shootings (-0.10%) and non-school mass shootings (-0.37%).

Lack of Access to Healthcare. The rationales of the ZSL model identifying articles disclosing a lack of access to mental/health care yielded the most topics. Our analysis of the topics produced 8 broad categories of accessing healthcare in which individual topics were placed: Assisted Suicide, Complications of Care, Denial of Care, Ineffective Care, Lack of Access to Physical Healthcare, Lack of Access to Mental Healthcare, and Management, Care, Treatment Alternatives, and Natural Disaster Loss-Related Stress. As it relates to management, care, and treatment alternatives, national publishers had higher probabilities of disclosing issues of substance use (0.21%) or technology and AI (0.53%) as alternatives to traditional care sought by many. Local publishers instead had a higher probability of disclosing when the alternative to care was social acceptance and freedom of self-expression (-0.64%) – which is less politically situated. Lack of access to mental healthcare had higher probabilities of being disclosed by local publishers reporting on access issues for specific populations with disparities in mental health or suicide related behavior: veterans (-0.86%), adolescents (-2.34%), and indigenous (-0.70%). And national publishers disclose when it relates to less easily enumerated contexts of care access issues like poverty (0.18%) or the burden of long-COVID (1.27%). National publishers disclose around policy related controversy when care is denied, abortion (0.09%) and LGBTQ+ Policy (0.85%).

Patterns of suppression in news communities. We now look at amplification and suppression among news communities by studying patterns of copying suicide-related articles. The largest number of suicide-related articles copied across news outlets occurred within community 2 (colored in lavender in Figure 4) with 2,209 articles copied. This community contained of a mix of national and local outlets with mostly centrist political leanings (i.e., Associated Press, ABC News, The Seattle Times, etc.) and was by far the largest.

Of the smaller communities that amplify suicide-related articles, many are communities of right-leaning and fringe outlets (as supported in Figure 5a). For example, community 12 (colored in green in Figure 4), copied the second most suicide-related articles, although magnitudes less than community 2 with only 157 articles copied. This community contains far-right and fringe outlets such as Infowars, RedState, and Veterans Today. Similarly, community 6 (colored in light blue in 4) copied 94 suicide-related articles and contained outlets like Fox News, Breitbart, and The Gateway Pundit.



(a) (b)

Figure 4: (a) Content sharing network, (b) number of suicide-related articles copied in each community.

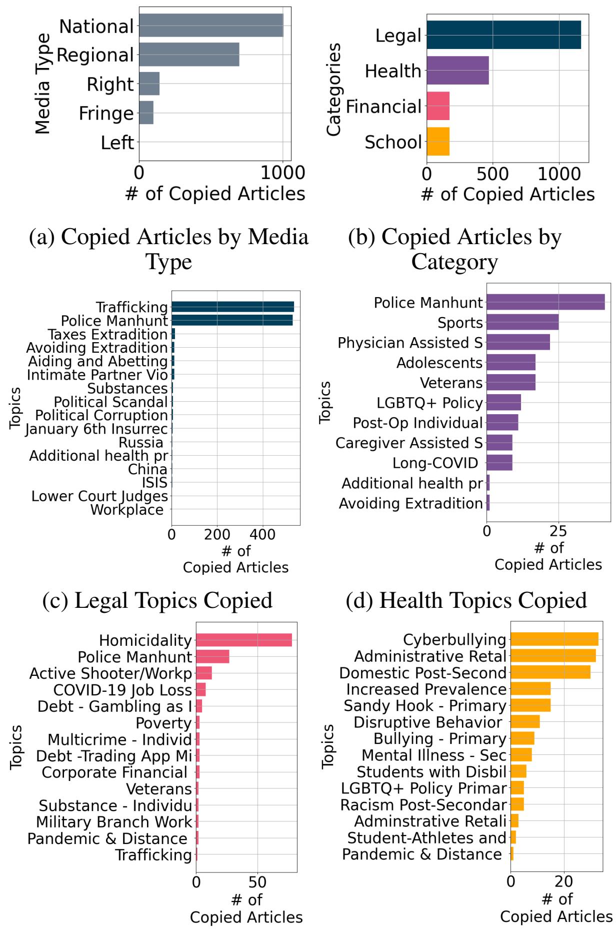
Community	Example Outlets
2	abcnews, usnews, theseattletimes
12	theepochtimes, veteranstoday, infowars
9	charlotteobserver, csmonitor, thehill
6	thegatewaypundit, breitbart, foxnews
4	rawstory, democracynow
25	latimes, dailyleader, valleytimesnews
7	newyorkpost, theconservativetreehouse
13	dailybeast, interpretermag

Table 4: Example news outlets in each community of the network shown in Figure 4a.

In contrast, left-leaning outlets copied only 3 articles, across all communities. This reflects the findings of Horne, Nørregaard, and Adalı (2019), which showed that content sharing occurs more frequently among right-leaning and fringe outlets than left-leaning outlets.

These differences across communities are in part due to both the size of the community and the types of outlets within each community outlets. However, when examining the proportion of suicide-related articles copied to total articles copied per community, we see that community 2 indeed copied proportionally more suicide-related articles than other communities. Specifically, 1.23% of all copied articles in community 2 were suicide-related articles, while all other communities copied proportionally less (ranging from 0.13% to 1.03% with an average 0.63% and median 0.35%). When examining all articles produced not just articles copied, the proportion of suicide-related articles produced to total articles produced per community remained small across all communities, with an average of 0.32% (median 0.29%) of articles produced within a community being suicide related.

When digging into the specific articles and topics copied, we see that the vast majority of copied articles are from two sub-topics within the Legal structural factor: Trafficking and Police Manhunt. Copied articles grouped under Trafficking were almost entirely coverage of Jeffrey Epstein's death and the trial of Ghislaine Maxwell. Coverage of Jeffrey Epstein's death crossed news communities, with coverage by national media, local media, and fringe media. This coverage ranged from facts of his death and the related trial to conspiracy theories. On the other hand, copied articles grouped under



(a) Copied Articles by Media Type

(b) Copied Articles by Category

(c) Legal Topics Copied

(d) Health Topics Copied

(e) Financial Topics Copied

(f) School Topics Copied

Figure 5: Suicide-related articles copied, broken down by (a) media type, (b) structural category, (c) topics within Legal category, (d) topics within Health category, (e) topics within Financial category, and (f) topics within School category.

Police Manhunt were much more diverse, including many individual stories ranging from national to local coverage. This subtopic was also commonly copied under Health and Financial (see Figures 5d and 5e).

Within the Health, Financial, and School structural factors, there was a wider variety of subtopics amplified, but these topics were copied magnitudes less than Trafficking and Police Manhunt within the Legal structural frame (Figures 5d-5f). The most frequently copied topic among the three structural factors contained 78 articles, while Trafficking and Police Manhunt within the Legal structural frame contained 536 and 531 articles, respectively.

Discussion

Theoretical Implications

Broadly, we found differences in disclosure, context, and copying of structural circumstances between news publishers by political leaning (left and right) and regionality (national and local). Still, these findings are descriptive, telling us little about suppression of these factors, without context.

Thus, we operationalize The Theory of Spiral of Silence (Noelle-Neumann 1974) as a theoretical framework of news media suppression of structural circumstances and implications for the suicide risk factor of social isolation. We discuss our findings through the three core mechanisms of suppression informed by the theory: *isolation*, *divergence*, and *suppression*.

Framing publisher categories within the NELA dataset allowed us to examine attribution of structural circumstances in their structural context. As public health scholars call out, existing work on suicide risk factors appear to ignore the overarching oppressive contexts in which people experience suicide behaviors and enable institutions to escape responsibility for adequately resourcing a response (Fitzpatrick 2018). By operationalizing the mechanism of **isolation** through theory, our findings not only affirm that these societal level considerations can be incorporated in suicide research, but that there are significant divergence disclosure between them – national and left leaning publishers disclosing legal circumstances more and right and local publishers disclosing financial, school, and health circumstances more. The mechanism of **divergence**, however, is essential to discerning the direction of structural context.

Classification of circumstances as performed in this paper through an LLM-based approach, enables us to compare disclosures. Topic modeling confirms the social contexts they represent – social and political determinants. Where statistical differences in disclosures appear counter-intuitive, between left and right, topics reveal that they only indicate the magnitude of the spiral. Left leaning publishers disclose to point out collectivist attitudes and systemic problems, while right leaning publishers blame the individual, aligns with partisan attitudes (Newport and Dugan 2017). This indicates that the potentially more isolating spiral of right leaning publishers is more pervasive than the communal spiral of the left – indicated by differences in frequency disclosures.

We offer that expanded use of communications theory in safe suicide reporting research will support findings that allow journalistic guidelines and prevention strategies to be more situated than general exclusion and inclusion criteria (SAVE 2020). This is especially valuable for furthering work exploring safe reporting and socio-demographic disparities in suicide – such as those of unsafe language and gender, noted by Foriest et al. (2024).

These patterns in reporting creating the spiral mechanisms of isolation and divergence, culminate in the third mechanism of **suppression.** Right leaning and left leaning publishers may be particularly motivated to suppress particular views and foster increased polarization explained through political typology (Iyengar and Hahn 2009). The patterns we detect are consistent with documented partisan positions and regional emphasis (Levendusky 2013). In the case of racial injustice and financial policy regarding taxes, left leaning respondents consistent opt for systemic and institutional change while right leaning respondents suggest little to nothing should be done in response to these social issues (Doherty et al. 2021). The same holds for immigration, health, and safety issues (Doherty et al. 2021). We find that topical suppression highlights the concerns of ethics scholars like Fitzpatrick. Attribution of the structural to individual and away from the political and institutional is likely to constrain the effectiveness of suicide prevention efforts and deflect institutional accountability (Fitzpatrick 2018). This is a potential motivator by political leaning and regionality as institutions exist on both levels.

Classification & Content Sharing For Prevention

Just as use of theory to center structural and political context can be insightful, so can the computational methods we have employed. With the extracted content sharing networks, we put suicide research and computational social science research in conversation in a novel way. Where most existing work has focused on social media (Kumar et al. 2015a; Sawhney et al. 2021; Singh et al. 2024), we bridge these methods with news media.

We find that copying of news reports of suicide with structural circumstance is consistent with prior work – increased sharing among far-right and fringe outlets (Horne, Nørregaard, and Adalı 2019). The types of stories most likely to be shared, by topic, are the most sensationalized or breaking news stories (Horne, Nørregaard, and Adalı 2019). This suggests that much of the knowledge of news network behavior holds true for suicide disclosures. However, our work further reveals that integrating these methods into a sensitive context requires multi-disciplinary considerations from practitioners. Prior work on news networks has yet to engage in the implications of news on adverse public health outcomes. The Spiral of Silence (Noelle-Neumann 1974) theory enables us to think about how established content sharing behaviors mediate suppression and unsafe reporting.

Given this, copying feels like a double-edged sword in the context of suicide and suicide prevention in news. The intuitive idea is that recommending content sharing to journalists when reporting on suicide could serve as a prevention or safe reporting strategy. Copying of articles with proper topical context, by left leaning and local publishers, of suicide deaths would theoretically help combat one mechanism of isolation just by making that perspective more prevalent in news. However, the articles most likely to get copied are likely sensationalized, problematic, and specific to one social environment. So copying just amplifies existing spirals and maybe even exacerbates suicide contagion (Niederkrotenthaler et al. 2020). Still, there are many nuances of copying behaviors that are yet to be studied, that constitute promising directions for future research.

Limitations and Future Work

Despite the contributions made in this paper, our study is not without limitations. We acknowledge the absence of important nuance in the ZSL learning prompts used. In suicidology and public health, single suicide deaths and homicide suicides are generally studied differently (Pirkis and Blood 2001). Additionally, several circumstances of suicide have varying definitions across each field of study. While our engineering of the ZSL prompt was informed by a validated annotation framework and definition for the selected circumstances (Chancellor et al. 2021), this approach is meant to demonstrate feasibility. Future work will require more precise prompt engineering or LLM fine-tuning to optimize and align the utility of computational strategies with existing knowledge in the study and prevention of suicide.

Finally, while the patterns in copying presented in this paper draw important attention to the potential mechanisms of suicide suppression, and suicide contagion illustrated in prior related work (Niederkrotenthaler et al. 2020), we do not thoroughly investigate the rationale for the copying practices. Also, as this is the first work of its kind, the type of copying we study here is near-verbatim copying. Future work should establish an expanded understanding of "copying" for more inclusive analysis.

Next, we recognize that suicide may be reported on proportionately less than other issues of injury and violence, such as violence against women (Mittal et al. 2024). An additional influence on reporting proportions is in the case of celebrity deaths. We note that our sample size of more than $9K$ articles, spanning diverse news publishers, and spanning over four years likely minimizes any effect that could potentially have a statistical sway in our findings. Still, we signal that future work should examine this potential relationship as a parallel to existing work highlighting the Werther/Papageno Effect of suicide contagion and celebrity suicides studied in Suicidology (Phillips 1974; Pirkis 2009; Niederkrotenthaler et al. 2010; Till and Niederkrotenthaler 2019) and recently in the context of web and social media data (Kumar et al. 2015b; Yuan et al. 2023). This leaves open questions about the other ways in which news reporting practices like copying also differ and defers further inferences to important future work on this topic.

Conclusion

We examined an essential component of worsening incidence suicide behaviors and deaths mediated by patterns in news publishing – social isolation. Using the Theory of Spiral of Silence, we operationalized the mechanisms of news mediated isolation into three conditions: isolation, divergence, and suppression. Further, we incorporated three important considerations for suicide prevention and research: unsafe reporting, isolation and inhibited help-seeking, and attention to structural determinants of suicide. We leveraged computational tools: zero-shot learning classification, BERTopic models, and extraction of content sharing networks to reveal differences in structural circumstance disclosure, context, and sharing of 22,021 news reports of suicide by political leaning and regionality. Left leaning publishers disclose structural factors with more collectivist and systemic contexts compared to right leaning publishers, whodisclose with individualist and personal culpability context. Patterns in content sharing reveal an opportunity for prevention professional to refine safe reporting guidelines to overcome mechanisms of suppression.

Acknowledgments

Foriest, Jain, and De Choudhury were partly supported through NIH grants R01 MH117172 and P50 MH115838, and a grant from the American Foundation for Suicide Prevention. Horne was funded by the CCI Board of Visitors Professorship. The findings of this paper do not represent the views or positions of any of the sponsors.

References

Bastian, M.; Heymann, S.; and Jacomy, M. 2009. Gephi: an open source software for exploring and manipulating networks. In *Proc. ICWSM*, volume 3, 361–362.

Brown, T.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J. D.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems*, 33: 1877–1901.

Castillo, E.; Hansen, H.; and Rocha, E. 2018. Position Statement on Mental Health Equity and the Social and Structural Determinants of Mental Health. *American Psychiatric Association*.

Centers for Disease Control and Prevention. 2023a. Disparities in Suicide. https://www.cdc.gov/suicide/facts/disparities-in-suicide.html. Last accessed: May 15, 2024.

Centers for Disease Control and Prevention. 2023b. Injury Center Research Priorities. https://www.cdc.gov/injury/researchpriorities/index.html. Last accessed: May 15, 2024.

Chancellor, S.; Sumner, S. A.; David-Ferdon, C.; Ahmad, T.; and De Choudhury, M. 2021. Suicide risk and protective factors in online support forum posts: annotation scheme development and validation study. *JMIR mental health*, 8(11): e24471.

Chandler, A.; Cover, R.; and Fitzpatrick, S. J. 2022. Critical suicide studies, between methodology and ethics: Introduction. *Health*, 26(1): 3–9.

Chen, T.; and Roberts, K. 2021. Negative life events and suicide in the national violent death reporting system. *Archives of suicide research*, 25(2): 238–252.

Chim, J.; Tsakalidis, A.; Gkoumas, D.; Atzil-Slonim, D.; Ophir, Y.; Zirikly, A.; Resnik, P.; and Liakata, M. 2024. Overview of the clpsych 2024 shared task: Leveraging large language models to identify evidence of suicidality risk in online posts. In *Proceedings of the 9th Workshop on Computational Linguistics and Clinical Psychology (CLPsych 2024)*, 177–190.

Choi, D.; Sumner, S. A.; Holland, K. M.; Draper, J.; Murphy, S.; Bowen, D. A.; Zwald, M.; Wang, J.; Law, R.; Taylor, J.; et al. 2020. Development of a machine learning model using multiple, heterogeneous data sources to estimate weekly US suicide fatalities. *JAMA network open*, 3(12): e2030932–e2030932.

Choudhury, M. D.; and Kıcıman, E. 2017. The Language of Social Support in Social Media and Its Effect on Suicidal Ideation Risk. *Proceedings of the ... International AAAI Conference on Weblogs and Social Media. International AAAI Conference on Weblogs and Social Media*, 2017: 32–41.

Coppersmith, G.; Leary, R.; Crutchley, P.; and Fine, A. 2018. Natural language processing of social media as screening for suicide risk. *Biomedical informatics insights*, 10: 1178222618792860.

Dawson, A.; and Silva, D. 2009. Suicide prevention: a task for public health and a role for public health ethics. *Journal of Public Mental Health*, 8(3): 4–6.

Doherty, C.; Kiley, J.; Asheer, N.; and Jordan, C. 2021. Beyond Red vs. Blue: The Political Typology. Technical report, Pew Research Center.

Durkheim, E. 2005. *Suicide: A study in sociology*. Routledge.

Fitzpatrick, S. J. 2018. Reshaping the ethics of suicide prevention: responsibility, inequality and action on the social determinants of suicide. *Public Health Ethics*, 11(2): 179–190.

Flek, L. 2020. Returning the N to NLP: Towards contextually personalized classification models. In *Proc. ACL*, 7828–7838.

Foriest, J. C.; Mittal, S.; Kim, E.; Carmichael, A.; Lennon, N.; Sumner, S. A.; and De Choudhury, M. 2024. News media framing of suicide circumstances and gender: mixed methods analysis. *JMIR mental health*, 11(1): e49879.

Gollapalli, S. D.; Zagatti, G. A.; and Ng, S. K. 2021. Suicide risk prediction by tracking self-harm aspects in tweets: Nus-ids at the clpsych 2021 shared task. In *Proc. CLPsych*, 93–98.

Grootendorst, M. R. 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *ArXiv*, abs/2203.05794.

Holman, M. S.; and Williams, M. N. 2022. Suicide risk and protective factors: a network approach. *Archives of suicide research*, 26(1): 137–154.

Horne, B. D.; Gruppi, M.; Joseph, K.; Green, J.; Wihbey, J. P.; and Adalı, S. 2022. NELA-Local: A dataset of US local news articles for the study of county-level news ecosystems. In *Proc. ICWSM*, volume 16, 1275–1284.

Horne, B. D.; Nørregaard, J.; and Adalı, S. 2019. Different spirals of sameness: A study of content sharing in mainstream and alternative media. In *Proc. ICWSM*, volume 13, 257–266.

Iyengar, S.; and Hahn, K. S. 2009. Red media, blue media: Evidence of ideological selectivity in media use. *Journal of communication*, 59(1): 19–39.

Jiang, A. Q.; Sablayrolles, A.; Mensch, A.; Bamford, C.; Chaplot, D. S.; de las Casas, D.; Bressand, F.; Lengyel, G.; Lample, G.; Saulnier, L.; Lavaud, L. R.; Lachaux, M.-A.; Stock, P.; Scao, T. L.; Lavril, T.; Wang, T.; Lacroix, T.; and Sayed, W. E. 2023. Mistral 7B. arXiv:2310.06825.

Jiang, A. Q.; Sablayrolles, A.; Roux, A.; Mensch, A.; Savary, B.; Bamford, C.; Chaplot, D. S.; Casas, D. d. l.; Hanna, E. B.; Bressand, F.; et al. 2024. Mixtral of experts. *arXiv preprint arXiv:2401.04088*.

Joiner, T. 2005. *Why people die by suicide*. Harvard University Press.

Kojima, T.; Gu, S. S.; Reid, M.; Matsuo, Y.; and Iwasawa, Y. 2022. Large Language Models are Zero-Shot Reasoners. In *Proc. Neurips*.

Kruskal, J. B. 1956. J., 1956. On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical Society*, 48–50.

Kumar, M.; Dredze, M.; Coppersmith, G.; and De Choudhury, M. 2015a. Detecting Changes in Suicide Content Manifested in Social Media Following Celebrity Suicides. In *Proc. HT*, HT '15, 85–94. ISBN 9781450333955.

Kumar, M.; Dredze, M.; Coppersmith, G.; and De Choudhury, M. 2015b. Detecting changes in suicide content manifested in social media following celebrity suicides. In *Proceedings of the 26th ACM conference on Hypertext & Social Media*, 85–94.

Levendusky, M. S. 2013. Why do partisan media polarize viewers? *American journal of political science*, 57(3): 611–623.

Liu, P.; Yuan, W.; Fu, J.; Jiang, Z.; Hayashi, H.; and Neubig, G. 2023. Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. *ACM Comput. Surv.*, 55(9).

McTernan, N.; Spillane, A.; Cully, G.; Cusack, E.; O'Reilly, T.; and Arensman, E. 2018. Media reporting of suicide and adherence to media guidelines. *International Journal of social psychiatry*, 64(6): 536–544.

Mittal, S.; Foriest, J. C.; Horne, B. D.; and De Choudhury, M. 2024. News Media and Violence Against Women: Understanding Framings of Stigma. In *Proc. ICWSM*, volume 18.

NCIPC, C. 2022. Suicide Prevention: Risk and Protective Factors. Suicide Prevention: Risk and Protective Factors. Centers for Disease Control and Prevention, National Center for Injury Prevention and Control. Accessed: 2018-12-06. https://www.cdc.gov/suicide/factors/index.htm.

Newport, F.; and Dugan, A. 2017. Partisan Differences Growing on a Number of Issues. *Gallup*. Last accessed May 15, 2024.

Niederkrotenthaler, T.; Braun, M.; Pirkis, J.; Till, B.; Stack, S.; Sinyor, M.; Tran, U. S.; Voracek, M.; Cheng, Q.; Arendt, F.; Scherr, S.; Yip, P. S. F.; and Spittal, M. J. 2020. Association between suicide reporting in the media and suicide: systematic review and meta-analysis. *The BMJ*, 368.

Niederkrotenthaler, T.; Voracek, M.; Herberth, A.; Till, B.; Strauss, M.; Etzersdorfer, E.; Eisenwort, B.; and Sonneck, G. 2010. Role of media reports in completed and prevented suicide: Werther v. Papageno effects. *British Journal of Psych*, 197(3): 234–243.

Noelle-Neumann, E. 1974. The spiral of silence a theory of public opinion. *Journal of communication*, 24(2): 43–51.

Nørregaard, J.; Horne, B. D.; and Adalı, S. 2019. NELA-GT-2018: A large multi-labelled news dataset for the study of misinformation in news articles. In *Proc. ICWSM*, volume 13, 630–638.

NVSS. 2024. Mortality 2018-2022 on CDC WONDER Online Database.

Oghaz, T. A.; Mutlu, E. Ç.; Jasser, J.; Yousefi, N.; and Garibay, I. 2020. Probabilistic model of narratives over topical trends in social media: A discrete time model. In *Proceedings of the 31st ACM Conference on Hypertext and Social Media*, 281–290.

Phillips, D. P. 1974. The influence of suggestion on suicide: Substantive and theoretical implications of the Werther effect. *American sociological review*, 340–354.

Pirkis, J. 2009. Suicide and the media. *Psychiatry*, 8(7): 269–271.

Pirkis, J.; and Blood, R. W. 2001. Suicide and the media: Part I. Reportage in nonfictional media. *Crisis: The Journal of Crisis Intervention and Suicide Prevention*, 22(4): 146.

Pirkis, J.; Blood, R. W.; Beautrais, A.; Burgess, P.; and Skehan, J. 2006. Media guidelines on the reporting of suicide. *Crisis*, 27(2): 82–87.

Roth, R.; Abraham, J.; Zinzow, H.; Wisniewski, P.; Khasawneh, A.; and Chalil Madathil, K. 2020. Evaluating news media reports on the'blue whale challenge'for adherence to suicide prevention safe messaging guidelines. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW1): 1–27.

Sainsbury, P. 1955. *Suicide in London*, volume 1. London.

SAMHSA. 2021. "Key substance use and mental health indicators in the United States: results from the 2022 National Survey on Drug Use and Health".

SAVE. 2020. Recommendations for reporting on Suicide. *Reporting on Suicide*. SAVE. Recommendations for reporting on Suicide. Reporting on Suicide. https://reportingonsuicide.org/recommendations/. Published 2020. Accessed October 12, 2022.

Sawhney, R.; Joshi, H.; Nobles, A.; and Shah, R. R. 2021. Towards Emotion- and Time-Aware Classification of Tweets to Assist Human Moderation for Suicide Prevention. *Proc. ICWSM*, 15(1): 609–620.

Singh, L. G.; Mao, J.; Mutalik, R.; and Middleton, S. 2024. Extracting and Summarizing Evidence of Suicidal Ideation in Social Media Contents Using Large Language Models. In *Proceedings of the 9th Workshop on Computational Linguistics and Clinical Psychology (CLPsych 2024)*, 218–226.

Sinyor, M.; Tran, U. S.; Garcia, D.; Till, B.; Voracek, M.; and Niederkrotenthaler, T. 2021. Suicide mortality in the United States following the suicides of Kate Spade and Anthony Bourdain. *Australian & New Zealand Journal of Psychiatry*, 55(6): 613–619.

Snapcrack. 2021. All the News Dataset. Accessed: 2024-05-15.

Song, K.; Tan, X.; Qin, T.; Lu, J.; and Liu, T.-Y. 2020. Mpnet: Masked and permuted pre-training for language understanding. *Advances in neural information processing systems*, 33: 16857–16867.

Starbird, K.; Arif, A.; Wilson, T.; Van Koevering, K.; Yefimova, K.; and Scarnecchia, D. 2018. Ecosystem or echo-system? exploring content sharing across alternative media domains. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 12.

Sumner, S. A.; Burke, M.; and Kooti, F. 2020. Adherence to suicide reporting guidelines by news shared on a social networking platform. *Proceedings of the National Academy of Sciences*, 117(28): 16267–16272.

Till, B.; and Niederkrotenthaler, T. 2019. Media and suicide: current state of research on Werther and Papageno effect—a literature review. In *Psychotherapie Forum*, volume 23, 120–128. Springer.

Touvron, H.; Martin, L.; Stone, K.; Albert, P.; Almahairi, A.; Babaei, Y.; Bashlykov, N.; Batra, S.; Bhargava, P.; Bhosale, S.; et al. 2023. Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*.

Wilkins, N.; Myers, L.; Kuehl, T.; Bauman, A.; and Hertz, M. 2018. Connecting the dots: state health department approaches to addressing shared risk and protective factors across multiple forms of violence. *Journal of public health management and practice*, 24: S32–S41.

Yuan, Y.; Saha, K.; Keller, B.; Isometsä, E. T.; and Aledavood, T. 2023. Mental health coping stories on social media: a causal-inference study of Papageno effect. In *Proceedings of the ACM Web Conference 2023*, 2677–2685.

Paper Checklist

1. For most authors...
   (a) Would answering this research question advance science without violating social contracts, such as violating privacy norms, perpetuating unfair profiling, exacerbating the socio-economic divide, or implying disrespect to societies or cultures? Yes, this work emphasizes the importance of considering social and political dynamics to understand harmful reporting practices. This work is rooted in expanding suicide prevention practices, methodologies, and intersections with computational social science. This work is intended as a bridge. This is described both in the related works and the discussion.
   (b) Do your main claims in the abstract and introduction accurately reflect the paper's contributions and scope? Yes, in the abstract, we succinctly characterize our context, findings, and methods. We mention all axes of analysis along with a novel theoretical framework application.
   (c) Do you clarify how the proposed methodological approach is appropriate for the claims made? Yes. Our methods and rationales are available in the methods section of each research question in the body of the paper. We align the methodological choices with existing applications of those methods in prior work either by task, such as classification, or by data source, such as the NELA dataset.
   (d) Do you clarify what are possible artifacts in the data used, given population-specific distributions? Yes. The section titled "Data" describes the characteristics of our data: source, number of articles, time period, validation data set, and meta information.
   (e) Did you describe the limitations of your work? Yes, we characterize the limitations, and practical considerations, of our work in the discussion and limitations sections.
   (f) Did you discuss any potential negative societal impacts of your work? Yes, we caveat in the Discussion and Limitations sections that while we think our approach could be very useful for multi-disciplinary work in suicide prevention and safe reporting, the subject matter for this work is sensitive. This requires extensive forethought and additional future work in the area.
   (g) Did you discuss any potential misuse of your work? No, while we consider it, we do not discuss any misuse of this work. As the methods of this work reveal existing problematic practices, rather than developing any tools to be deployed in any setting, this work is unlikely to be misused.
   (h) Did you describe steps taken to prevent or mitigate potential negative outcomes of the research, such as data and model documentation, data anonymization, responsible release, access control, and the reproducibility of findings? Yes, in the ethics statement preceding this checklist, we describe out multi-disciplinary expertise and integration of best practices in research. We worked with deidentified data. We also use paraphrased rationales from the LLM in lieu of graphic excerpts from news articles in the paper to reduce traceability.
   (i) Have you read the ethics review guidelines and ensured that your paper conforms to them? Yes, we reviewed the ethics review guidelines provided at https://aaai.org/about-aaai/ethics-and-diversity/ and are completing the ethics checklist.
2. Additionally, if your study involves hypotheses testing...
   (a) Did you clearly state the assumptions underlying all theoretical results? NA
   (b) Have you provided justifications for all theoretical results? NA
   (c) Did you discuss competing hypotheses or theories that might challenge or complement your theoretical results? NA
   (d) Have you considered alternative mechanisms or explanations that might account for the same outcomes observed in your study? NA
   (e) Did you address potential biases or limitations in your theoretical framework? NA
   (f) Have you related your theoretical results to the existing literature in social science? NA
   (g) Did you discuss the implications of your theoretical results for policy, practice, or further research in the social science domain? NA
3. Additionally, if you are including theoretical proofs...
   (a) Did you state the full set of assumptions of all theoretical results? ANA
   (b) Did you include complete proofs of all theoretical results? NA
4. Additionally, if you ran machine learning experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? Yes, we provide a link to all code, data, and instructions for reproducibility in the conclusion of this paper.
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? NA
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? NA
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? NA
   (e) Do you justify how the proposed evaluation is sufficient and appropriate to the claims made? Yes, we provide rationale for evaluating model performance in the methods sections of each research question in which ML approaches were used.
   (f) Do you discuss what is "the cost" of misclassification and fault (in)tolerance? Yes, we discuss that in our Error Analysis sections as well as the Limitations and Future Work section.
5. Additionally, if you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? Yes, we thoroughly cite the creators of both

datasets and their respective models in each research question's methods section.

(b) Did you mention the license of the assets? NA, assets are not licensed

(c) Did you include any new assets in the supplemental material or as a URL? Answer

(d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? NA, datasets used in this text are all open source and for public use.

(e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? Yes, we provide a content warning related to the sensitivity of the dataset and related excerpts covered in the paper. This is located below the abstract

(f) If you are curating or releasing new datasets, did you discuss how you intend to make your datasets FAIR (see **?**)? NA

(g) If you are curating or releasing new datasets, did you create a Datasheet for the Dataset (see **?**)? NA

6. Additionally, if you used crowdsourcing or conducted research with human subjects...

(a) Did you include the full text of instructions given to participants and screenshots? NA

(b) Did you describe any potential participant risks, with mentions of Institutional Review Board (IRB) approvals? NA

(c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? NA

(d) Did you discuss how data is stored, shared, and deidentified? NA

Ethical Statement

The authors of this work have professional expertise in computational social science, public and mental health, news media research, and injury and violence prevention. Thus we recognize that this approach has implications for those experiencing suicide behaviors. The analysis requires the collapsing of individual stories of people's real lived experiences and deaths into aggregate and computationally discernible patterns of reporting. And even though these methods were performed mindfully, illustrated in careful curation of topic labels intended to minimize or reinforce use of stigmatizing language, there are complexities of circumstances of suicide that are subsumed by our lenses of structural factors and news communities. We offer that future work of computational social scientists and public health practitioners leveraging these lenses or methods carefully consider this reality as well when developing tools to address suppression of suicide circumstances, or any issue of injury and violence, in news.

Appendix

Relative Comparison of LLMs in Classification of Structural Factors

Based on the observations in Table 2 of the paper, GPT-3.5 outperforms GPT-4 in specific cases, such as identifying financial problems, despite GPT-4 generally performing better overall. This could be due to several factors:

1. **Model Size and Complexity**: GPT-4, being a more complex model with significantly more parameters, might overfit certain categories, especially when the training data is limited or the classification task requires high generalization. In contrast, GPT-3.5 may perform better due to its simpler architecture, which helps it generalize better in specific cases like financial problems where the patterns are more straightforward.
2. **Task Specialization**: In the financial problem category, GPT-3.5 shows a higher accuracy and AUC compared to GPT-4, which could indicate that GPT-3.5's model tuning or pretraining allows it to capture certain features more effectively for this task. This may be because the financial problems could involve less complex, more concrete patterns that GPT-3.5 captures more efficiently, while GPT-4, with its broader scope, might incorporate unnecessary complexity.
3. **Prompt Sensitivity**: The prompts used for zero-shot learning might have been better suited for GPT-3.5 in certain categories, as observed with financial and legal problems. This difference in prompt sensitivity could explain why GPT-3.5 outperforms GPT-4 in identifying financial issues, even though GPT-4 generally excels in other areas.

In addition to GPT-3.5 and GPT-4, the performance of other large language models (LLMs) such as Mixtral, Mistral, and LLaMa is also noteworthy when evaluated across the four structural factors: financial/job problems, legal problems, school/academic problems, and lack of access to health/mental healthcare. While GPT models generally lead in most categories, Mixtral and Mistral exhibit some promising results in specific contexts. For example, in the legal problem category, Mixtral achieved an accuracy of 0.75, and Mistral had an accuracy of 0.76, suggesting that these models are particularly effective in detecting legal factors, though they fall short of the higher accuracy seen in GPT models.

For the school/academic problem factor, Mistral performed notably well with an accuracy of 0.96 using zero-shot prompting, which is close to GPT-4's performance (0.98). This indicates that Mistral can successfully capture the nuances of school-related issues, likely due to its capacity to handle context-rich, detailed patterns in textual data. Mixtral, however, struggled in this category, showing lower precision and recall, likely due to its smaller parameter size and perhaps less effective tuning for tasks requiring deep contextual understanding.

On the other hand, in detecting issues related to lack of access to health/mental healthcare, Mixtral and LLaMa performed below expectations, with Mixtral achieving an accuracy of 0.65 and LLaMa trailing further behind. This underperformance could be attributed to these models' reduced capacity to handle the more abstract and multifaceted nature of healthcare access issues, which might require the model to understand implicit social and systemic context, something that larger models like GPT-4 handle better. The complexity of healthcare-related structural factors, especially when co-occurring with other factors, likely presents a challenge for smaller models such as Mixtral and LLaMa.

Overall, while GPT-3.5 and GPT-4 lead in the majority of the structural factors, Mixtral and Mistral can be viable alternatives in certain specialized contexts, particularly where specific factor detection, like legal or school-related issues, is involved. However, their performance drops notably in categories requiring more comprehensive and nuanced understanding, such as healthcare access. The differences in performance across these models further highlight the need for choosing the right model depending on the task's complexity and context, especially in the context of sensitive issues like suicide-related circumstances.

LLM Classifier Performance With Respect to Human Annotations

As noted in Table 2, ZSL and FSL models were evaluated against a heldout 20% sample of the 200 human-labeled articles from the `All The News` dataset, to assess their performance in detecting structural factors related to suicide. In terms of raw accuracy and AUC, ZSL demonstrated commendable results across different structural factors. FSL which incorporated additional examples from the training data, performed similarly well, albeit with poorer metrics in certain categories.

Qualitatively, both ZSL and FSL models showed remarkable consistency with human-labeled annotations in simpler, more concrete structural factors, such as legal and financial problems. Human annotators of the `All The News` dataset identified these factors based on explicit textual cues (e.g., references to court cases, financial hardship), which the models were also able to recognize effectively. This suggests that for tasks with clearly defined, literal indicators, both ZSL and FSL can closely replicate human annotation quality.

However, in categories where the structural factors are more abstract or require deeper contextual understanding, such as lack of access to healthcare, the models exhibited occasional misclassifications compared to the human-labeled data. In several cases, the ZSL model inferred the presence of healthcare problems from descriptions of emotional turmoil or mental health struggles, even when the lack of access to treatment was not explicitly mentioned. Human annotators, by contrast, tended to make more nuanced distinctions, only labeling articles as relating to healthcare access when clear systemic failures were discussed. This qualitative difference in interpretation highlights a limitation in the models' ability to infer complex structural issues without sufficient explicit indicators in the text.

Overall, while both ZSL and FSL demonstrate strong alignment with human annotations for clearly defined factors, their qualitative performance falls short when handling more abstract or contextually rich categories.

Code —
https://github.com/NewsSuppressionSF/Suppression-SF

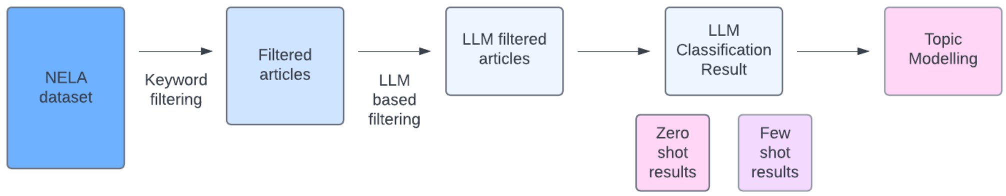


Figure A1: Overall methodology for filtering the data and classification

Suicide is the act of intentionally causing one's own death or taking their own life by any means. You are a public health professional. You are going through news articles from across the USA that mention suicide and your task is to identify the cause of the suicide talked about in the article. The article may not use the word 'suicide', instead explain it in another way, such as 'he took his own life' or a different expression that has the same meaning. You have to categorize the articles in one or more of the following categories: *1) 'Financial/Job Problem': If there is mention of a person or multiple people who killed themselves due to financial or work related problems or bad financial state or an incident at their place of employment.* *2) 'School Problem': If there is mention of a person or multiple people who killed themselves due to an academic or school related problem.* *3) 'Health Problem': If there is mention of a person or multiple people who killed themselves due to lack of access to mental or physical health care or treatment.* *4) 'Legal Problem': If there is mention of a person or multiple people who killed themselves after facing a legal problem or being involved or suspected or targeted in a violation of the law. 5) 'None': If there is no explicit mention of a person who committed suicide or killed themselves or if suicide was attempted or threatened but the person did not die or if the reason of suicide is not aligned with any of the above 4 categories.* *Now, read the following news article answer the following 2 questions: 1. Is there mention of a person or multiple people who died after committing suicide or killed themselves? Provide your answer in the following format: Answer 1: Yes or No 2. Identify which of the above categories does the article belong to. An article can belong to more than one category. Also provide a reasoning for the same in the following format:* *Answer 1:* *Answer 2:* *Reasoning:* *Article:*

Table A1: Zero shot prompt input template for classification using LLMs

Following are the examples of news articles talking about suicide and the answer to which category it falls into along with the reasoning behind that answer: *Article: Article 1* *Answer: Label* *Reasoning: Reason* *......* *Article: Article 4* *Answer: Label* *Reasoning: Reason*

Table A2: Additional text in FSL prompt, on top of the ZSL prompt, given in Table A1.

Topic Category	Subcategory	Systemic vs. Individual	Right vs. Left	National vs. Local
Crime	Substance - Individual Debt	Individual	-0.32%	-2.60%
Crime	Cybercrime - Individual Scammed by Phishing	Systemic	3.90%	-0.57%
Crime	Corporate Financial - Individual	Systemic	3.54%	0.99%
Crime	Corporate Financial - Alex Murdoch	Systemic	0.43%	1.06%
Crime	Multicrime - Individual Debt	Individual	0.43%	1.32%
Crime	Corporate Financial - Ponzi Schemes & Trafficking	Individual	-1.95%	1.43%
Crime	Cybercrime - Multiple Social Media Hackers	Individual	-2.30%	1.67%
Financial Risk Taking	Debt - Gambling as Income	Individual	-0.23%	-0.82%
Financial Risk Taking	Gambling Addiction	Systemic	1.82%	-0.73%
Financial Risk Taking	Debt -Trading App Misunderstanding	Individual	-2.08%	2.04%
Housing	Fear of Homelessness	Systemic	1.93%	-2.31%
Housing	Eviction	Systemic	1.04%	-0.24%
Housing	Foreclosure and Divorce	Systemic	0.26%	0.62%
Job Related Stress	Homicidality	Individual	-1.32%	-7.31%
Job Related Stress	Active Shooter/Workplace Violence	Individual	0.45%	0.57%
Job Related Stress	Military Branch Working Conditions	Individual	-0.43%	0.93%
Systemic Economic Hardship	Rural Farmers in Financial Distress	Systemic	-0.34%	0.27%
Systemic Economic Hardship	COVID-19 Job Loss	Systemic	-0.90%	1.30%
Systemic Economic Hardship	Rural Farmers in Financial Distress	Systemic	-2.29%	1.57%

Table A3: Topic modeling results for Financial/Job Problem

Topic Category	Subcategory	Systemic vs. Individual	Right vs. Left	National vs. Local
Bullying & Harassment	Cyberbullying - Physical Appearance	Systemic	2.63%	3.82%
Bullying & Harassment	LGBTQ+	Systemic	-0.55%	-1.37%
Bullying & Harassment	Cyberbullying - Sexuality	Systemic	1.98%	5.16%
Bullying & Harassment	Cyberbullying - Online Extortion	Systemic	0.20%	1.29%
Discrimination	LGBTQ+ Misgendering Post-Secondary	Systemic	4.13%	0.05%
Discrimination	LGBTQ+ Policy Primary & Secondary	Systemic	-0.69%	-1.04%
Discrimination	Racism Post-Secondary	Systemic	0.69%	-1.05%
Discrimination	Racism Primary	Systemic	-0.42%	-0.98%
Discrimination	Students with Disabilities	Systemic	-1.33%	-0.07%
Shooting	International Post-Secondary Shooting	Systemic	4.36%	0.33%
Shooting	Administrative Retaliation - Post-Secondary	Individual	-1.31%	0.93%
Shooting	Domestic Post-Secondary	Individual	-1.18%	-2.20%
Shooting	Sandy Hook - Primary	Individual	-1.09%	-2.13%
Shooting	Bullying - Primary	Individual	-0.94%	-0.77%
Shooting	Mental Illness - Secondary	Systemic	2.91%	-0.94%
Shooting	Disruptive Behavior - Secondary	Individual	-2.07%	1.53%
Shooting	Secondary	Individual	0.03%	-1.31%
Shooting	Summary of Many Mass Shootings - Secondary	Individual	-0.84%	-0.42%
Shooting	Administrative Retaliation - Secondary	Individual	0.91%	2.75%
Shooting	Administrative Retaliation - Secondary	Individual	-0.89%	-0.09%
Social Media	Exposure to Negative Content	Systemic	-2.65%	2.00%
Student Mental Health Crisis	Pandemic & Distance Learning	Systemic	4.42%	-0.24%
Student Mental Health Crisis	Increased Prevalence of Suicidality	Systemic	0.02%	-6.24%
Student Mental Health Crisis	Student-Athletes and Performance Pressure	Systemic	-1.25%	-0.55%

Table A4: Topic modeling results for School or Academic Problem

Topic Category	Subcategory	Systemic vs. Individual	Right vs. Left	National vs. Local
Cybercrime	Doxxing	Individual	-0.87%	0.73%
Evading Prosecution	Financial	Individual	-0.09%	0.19%
Evading Prosecution	Homicide	Individual	-0.15%	0.46%
Incarceration Related Stress	Avoiding Extradition	Individual	-0.02%	0.27%
Incarceration Related Stress	Alleged Homicide	Individual	-0.00%	-0.91%
Incarceration Related Stress	Solitary Confinement	Systemic	-0.38%	0.25%
Incarceration Related Stress	Additional health problems	Systemic	0.44%	-0.11%
Incarceration Related Stress	Homicide Execution	Individual	-0.02%	0.09%
Incarceration Related Stress	Juvenile with Self-Harm History	Systemic	-0.26%	-0.03%
Incarceration Related Stress	Taxes Extradition	Individual	0.53%	-0.29%
Incarceration Related Stress	War Crime Execution	Individual	-0.31%	-0.03%
Mass Shootings	Non-School	Individual	-0.20%	-0.37%
Mass Shootings	School	Individual	-0.23%	-0.10%
Pre-Arrest/Pre-Trial	Intimate Partner Violence	Systemic	1.01%	0.45%
Pre-Arrest/Pre-Trial	Police Manhunt	Systemic	1.68%	-4.40%
Pre-Arrest/Pre-Trial	Aiding and Abetting	Individual	-0.17%	-0.20%
Pre-Arrest/Pre-Trial	Political Corruption	Individual	-0.35%	0.14%
Pre-Arrest/Pre-Trial	Substances	Systemic	0.28%	-0.30%
Pre-Arrest/Pre-Trial	Embezzlement	Individual	-0.26%	-0.11%
Pre-Arrest/Pre-Trial	Serial Killer	Individual	0.13%	0.20%
Retaliation Over Legal Decisions	Lower Court Judges	Systemic	-0.42%	0.45%
Retaliation Over Legal Decisions	Federal Judge: Roe v. Wade	Systemic	0.31%	0.15%
Sex Crime	Child Pornography	Individual	-0.26%	0.25%
Sex Crime	Political Scandal	Individual	-0.37%	0.43%
Sex Crime	Trafficking	Individual	-1.96%	3.17%
Sex Crime	Workplace	Individual	-0.46%	0.43%
Terrorism	Revolutionary Terrorism	Individual	-0.11%	0.15%
Terrorism	January 6th Insurrection	Systemic	0.60%	-0.15%
Terrorism	Bomber	Individual	0.12%	0.26%
Terrorism	China	Systemic	0.37%	-0.13%
Terrorism	ISIS	Systemic	0.01%	0.16%
Terrorism	Russia	Individual	-0.09%	0.39%

Table A5: Topic modeling results for Legal Problem

Topic Category	Subcategory	Systemic vs. Individual	Right vs. Left	National vs. Local
Assisted Suicide	Physician Assisted Suicide	Systemic	1.07%	-0.07%
Assisted Suicide	Caregiver Assisted Suicide	Individual	-0.27%	0.88%
Complications of Care	Post-Op Individual	Systemic	0.10%	-0.16%
Complications of Care	Systemic Lack of Care for HIV	Systemic	-0.52%	0.44%
Denial of Care	Sports	Individual	-0.39%	0.48%
Denial of Care	Incarceration	Systemic	-0.06%	-0.44%
Denial of Care	Mental Health Care	Systemic	-0.01%	-0.47%
Denial of Care	Abortion Care	Systemic	-0.08%	0.09%
Denial of Care	LGBTQ+ Policy	Systemic	0.35%	0.85%
Ineffective Care	Mental Health Care	Systemic	-0.36%	0.10%
Lack of Access to Mental Health Care	Poverty	Systemic	-0.58%	0.18%
Lack of Access to Mental Health Care	Veterans	Systemic	1.75%	-0.86%
Lack of Access to Mental Health Care	Adolescents	Systemic	-1.09%	-2.34%
Lack of Access to Mental Health Care	Indigenous Communities	Systemic	0.01%	-0.70%
Lack of Access to Physical Healthcare	Long-COVID	Individual	-0.11%	1.27%
Lack of Access to Physical Healthcare	Long-COVID	Individual	-0.22%	0.32%
Legal Intervention	Mandated Care	Systemic	0.46%	-0.21%
Management, Care, and Treatment Alternatives	Self-Expression	Individual	0.00%	-0.64%
Management, Care, and Treatment Alternatives	Substances	Individual	-0.32%	0.21%
Management, Care, and Treatment Alternatives	Technology/AI	Individual	-0.63%	0.53%
Natural Disaster Loss-Related Stress	Hurricane	Systemic	0.10%	-0.33%

Table A6: Topic modeling results for Lack of Access to Health/Mental Health Care

Factor	Systemic/Institutional	95% CI	Individualism	95% CI
Financial/Job	-0.48	[-0.958, 0.005]	0.86	[0.200, 1.516]
Legal	-0.34	[-0.446, -0.234]	0.29	[0.137, 0.435]
School	-0.47	[-0.900, -0.036]	1.82	[1.168, 2.470]
Health	-0.12	[-0.253, 0.023]	0.42	[0.099, 0.743]

Table A7: Average percentage difference for topics interpreted as systemic/institutional vs. individual

	Financial (95% CI)	Legal (95% CI)	School (95% CI)	Health (95% CI)
Left	[0.028, 0.052]	[0.010, 0.034]	[0.023, 0.031]	[0.011, 0.018]
Right	[0.029, 0.052]	[0.010, 0.033]	[0.024, 0.035]	[0.010, 0.018]
Left vs. Right	[-0.326, 0.497]	[-0.151, 0.058]	[-0.124, 0.712]	[-0.169, 0.093]
National	[0.028, 0.052]	[0.010, 0.034]	[0.023, 0.032]	[0.010, 0.018]
Local	[0.024, 0.056]	[0.009, 0.034]	[0.025, 0.034]	[0.010, 0.019]
National vs. Local	[-0.543, 0.458]	[-0.135, 0.227]	[-0.521, 0.391]	[-0.208, 0.125]

Table A8: Average percentage difference for Financial, Legal, School, and Health factors across different categories