Long-Term Development of Attitudes towards Schizophrenia and Depression in Scientific Abstracts

Ivan Nenchev [1,2], Tatjana Scheffler [3], Lisa Raithel [4], Elif Kara [6], Benjamin Wilck [5], Maren Rabe [1], Philip Stötzner [1], Christiane Montag [1]

[1] Department of Psychiatry and Psychotherapy, Charité Campus Mitte, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health,

[2] Berlin Institute of Health at Charité – Universitätsmedizin Berlin,

[3] German Studies, Ruhr-Universität Bochum,

[4] Quality & Usability Lab, Technische Universität Berlin

[5] The Hebrew University of Jerusalem, Martin Buber Society of Fellows in the Humanities and Social Sciences

[6] Freie Universität Berlin

ivan.nenchev@charite.de

Abstract

We present a study investigating the linguistic sentiment associated with schizophrenia and depression in research-based texts. To this end, we construct a corpus of over 260,000 PubMed abstracts published between 1975 and 2025, covering both disorders. For sentiment analysis, we fine-tune two sentence-transformer models using SetFit with a training dataset consisting of sentences rated for valence by psychiatrists and clinical psychologists. Our analysis identifies significant temporal trends and differences between the two conditions. While the mean positive sentiment in abstracts and titles increases over time, a more detailed analysis reveals a marked rise in both maximum negative and maximum positive sentiment, suggesting a shift toward more polarized language. Notably, sentiment in abstracts on schizophrenia is significantly more negative overall. Furthermore, an exploratory analysis indicates that negative sentences are disproportionately concentrated at the beginning of abstracts. These findings suggest that linguistic style in scientific literature is evolving. We discuss the broader ethical and societal implications of these results and propose recommendations for more cautious language use in scientific discourse.

1 Introduction

According to the ICD-10, schizophrenic disorders (F20–F29) are defined by "fundamental and characteristic distortions of thinking and perception, and affects that are inappropriate or blunted," while recurrent depressive disorder is characterized by "repeated episodes of depression as described for depressive episode (F32.-), without any history of independent episodes of mood elevation and increased energy (mania)" (World Health Organization, 2016). Research articles on these conditions often begin with sentences like: *"Schizophrenia is among the most severe and debilitating of psychiatric disorders"* (Schultz and Andreasen, 1999), and *"Major depression is a common illness that severely limits psychosocial functioning and diminishes quality of life"* (Malhi and Mann, 2018). What is striking about such sentences is that they tend to convey a negative sentiment and pessimism. Importantly, neither schizophrenia nor depression are invariably linked to a poor prognosis. Although some people experience persistent symptoms, others have a more favorable course. In the case of schizophrenia for example, a subset of patients experiences only a single psychotic episode, followed by stable remission (Alvarez-Jimenez et al., 2011; Molstrom et al., 2022), while others achieve functional recovery with appropriate interventions, including pharmacological and psychosocial treatments (Phahladira et al., 2020). The prognosis is influenced by a variety of factors, including early intervention (Howes et al., 2021), adherence to treatment (Fang et al., 2022) and psychotherapy (Lysaker et al., 2010), social support, and individual resilience (Wambua et al., 2020).

Schizophrenia and depression are both classified as severe mental illnesses and are among the most extensively studied psychiatric conditions in biomedical research. Despite this com-

Proceedings of the Fourth Workshop on NLP for Positive Impact (NLP4PI), pages 99–113
July 31, 2025 ©2025 Association for Computational Linguistics

mon ground, emerging evidence suggests a divergence in societal perception: while public stigma surrounding depression has declined in recent decades—reflecting growing awareness and acceptance—schizophrenia continues to be associated with persistent or even increasing stigma (Pescosolido et al., 2021; Schomerus et al., 2022). The way schizophrenia and depression are described in the scientific literature may play a crucial role in shaping public and professional perceptions of these disorders. Different linguistic framings emphasize distinct aspects of the conditions, influencing attitudes toward prognosis, treatment, and stigma. The examples presented above highlight how variations in language can convey different sentiments about the nature and course of these disorders.

Scientific publications serve as a primary medium for disseminating objective knowledge about psychiatric conditions, yet the language used in these texts can shape both clinical practice and public discourse. Over time, shifts in linguistic style and sentiment within academic literature may reflect broader developments in scientific understanding, medical advancements, and societal attitudes toward mental illness. Recent advances in natural language processing (NLP) enable the analysis of large linguistic datasets in a systematic and replicable manner, allowing researchers to uncover patterns of linguistic representation that may not be immediately apparent in individual texts. In this study, we analyze the linguistic style of scientific publications over the past 50 years, with a particular focus on sentiment, to examine how perspectives on schizophrenia and depression have evolved.

2 Background

2.1 Sentiment analysis

Sentiment analysis is a natural language processing (NLP) technique used to determine the polarity of a text (for a recent review see Hartmann et al., 2023 and Wankhade et al., 2022). Early sentiment analysis techniques relied on dictionary-based approaches, where predefined lexicons assigned sentiment scores to words and aggregated them to determine the overall polarity of a text (Hutto and Gilbert, 2014; Tausczik and Pennebaker, 2010). Although these methods were interpretable and computationally efficient, they struggled with context-dependent sentiment, negation handling, and domain-specific language variations. Modern sentiment analysis models leverage deep learning and transformer-based architectures to improve accuracy across diverse contexts. One such modern approach is SetFit (Sentence Transformer Fine-tuning) (Tunstall et al., 2022), a few-shot learning technique that fine-tunes sentence embeddings for sentiment classification. Unlike traditional transformer-based models like BERT, SetFit requires significantly fewer labeled examples while maintaining high accuracy, which makes it particularly useful for domain-specific sentiment analysis with limited annotated data. Using contrastive learning during fine-tuning, SetFit enhances the quality of sentence representations, allowing for more nuanced sentiment estimation. We use a domain-specific model fine-tuned for this study because most existing models are trained on tweets or product reviews, making them unsuitable for evaluating scientific texts. Our goal is to capture sentiment within a highly specific domain—abstracts of articles on schizophrenia and depression. This approach is particularly beneficial for analyzing scientific texts, where sentiment is often subtle, context-dependent, and requires domain expertise to interpret accurately.

2.2 Sentiment analysis of medical texts

Over the past decade, sentiment analysis has been increasingly applied to scientific texts, consistently revealing a shift toward more positive language. Early studies relied on predefined dictionaries to track sentiment changes in PubMed abstracts. Vinkers et al. (2015) examined 50 predefined positive and negative terms and found a rise in both, a finding later corroborated by Cao et al. (2021), who expanded the analysis to 2.2 million articles and observed a stronger increase in positive wording. Wen and Lei (2022) extended this research across 12 disciplines, applying the R packages Syuzhet and Sentimentr to 775,000 abstracts. Similarly, Edlinger et al. (2023) used VADER sentiment analysis[1] on 2.3 million MEDLINE abstracts from psychology, biology, and physics, reporting that positive language became especially prevalent toward the end of abstracts. More recent studies have used deep learning techniques: Myszewski et al. (2022) fine-tuned a BioBERT classifier to analyze sentiment trends in human and veterinary medical trials, confirming the growing prevalence of positive lan-

[1] https://github.com/cjhutto/vaderSentiment

guage.

Despite these advancements, sentiment analysis in psychology and psychiatry remains underexplored. Baes et al. (2022) examined 829,701 psychology abstracts (1970–2017) using LIWC (a dictionary based approach Tausczik and Pennebaker, 2010, identifying an increase in both positive and negative sentiment over time. Perlis and Jones (2024) employed zero-shot learning with GPT-4 to analyze sentiment in 12,000 abstracts from high-impact medical journals (2017–2022), finding that psychiatry abstracts were rated as more negative and less positive compared to those in cardiology, oncology, and neurology. The authors question whether the negative sentiment in psychiatric texts reinforces negative attitudes toward psychiatry among medical practitioners and, potentially, the broader community. These studies underscore the need for a more nuanced analysis of sentiment trends in psychiatry and psychology, particularly using advanced NLP methods to investigate how sentiment is shaped by disciplinary conventions and publication practices.

2.3 Negativity and positivity bias

Negative and positive information are processed asymmetrically, giving rise to the so-called negativity and positivity biases. Negative information tends to capture more attention (Veerapa et al., 2020), is more deeply encoded in memory, and is recalled more easily than positive information (Williams et al., 2022). As a result, people assign greater weight to negative traits or behaviors when forming impressions of others. Additionally, individuals engage in more causal reasoning for negative events, seeking explanations for their occurrence. In contrast, positive information is processed more quickly as it is often linked to a greater number of cognitive associations, facilitating learning. While negativity bias promotes vigilance and caution, positivity bias enhances cognitive efficiency and supports adaptive behaviors. Together, these biases shape decision-making, social judgments, and memory processes in everyday life. Rozin and Royzman (2001) describe a so called negativity dominance — the tendency for combinations of negative and positive information to be evaluated more negatively than the sum of their individual subjective valences would predict. This suggests that negative information exert a disproportionate influence when mixed with positive one, ultimately skewing the overall impression toward the negative.

Unsurprisingly, negative stereotypes form much easily and are harder to change compared to positive ones (Baumeister et al., 2001). Paolini et al. (2024) conducted a meta-analysis on contact between groups, showing that while positive contact systematically reduces prejudice, negative contact has a significantly stronger effect in increasing it. This asymmetry reflects the negativity bias, where adverse interactions carry more weight than beneficial ones. Furthermore, negativity bias is stronger in interactions with stigmatized low-status outgroups, especially when stigma is not concealable, in informal and nonintimate settings, and within collectivistic societies [2]. Furthermore, Bellucci (2023) show that the sequence of presenting positive and negative information affects recall, with negatively valenced information being more likely remembered when it precedes positively valenced information.

This study analyzes the sentiment of scientific abstracts on schizophrenia and depression using domain-specific models fine-tuned for this purpose. To our knowledge, this is the first attempt within our field to move beyond dictionary-based methods and pretrained models, providing a more nuanced assessment.

3 Materials and Methods

In the following section, we describe the PubMed corpus construction, the fine-tuning procedure for sentiment estimation, and the statistical analysis of the results.

3.1 PubMed corpus

We compiled a corpus of 282,666 abstracts from scientific publications published between 1975 and 2025 using Biopython and Entrez, with the query: ((Schizophrenia[MeSH Terms] OR Depression[MeSH Terms]) OR (Schizophrenia[Title] OR Depression[Title])). Each abstract was assigned to its corresponding publication year. Abstracts without a recorded publication year (n=20,032) were excluded from further analysis. The remaining abstracts were categorized into three groups: "schizophrenia," "depression," or "schizophrenia and depression." Abstracts classified as addressing both conditions (n=2077) were excluded from

[2] In collectivistic societies, group harmony and social cohesion are prioritized, contrasting with individualistic societies where autonomy and individual goals take precedence. The individualism-collectivism dichotomy is a fundamental dimension of cultural diversity.

subsequent analyses. The remaining abstracts (n=260,557) were segmented into individual sentences using spaCy (Montani et al., 2023). Table 5 in the Appendix provides an overview of the characteristics of the corpus. The corpus exhibits temporal imbalance characteristic of actual scientific publishing patterns, with substantially fewer abstracts in earlier years compared to recent decades, reflecting the exponential growth in biomedical research output over time (see Statistical Analysis section for methodological adjustments addressing this imbalance).

3.2 Finetung the SetFit models and sentiment extraction

To create a training dataset for fine-tuning the SetFit model, we used 12 abstracts from the *Lancet* Seminar series on schizophrenia (n=5) and depression (n=7), published between 1999 and 2022. Although the dataset is relatively small, it includes *all* available *Lancet* Seminar abstracts on these conditions. This article type, published only every few years, summarizes recent scientific advances and is authored by leading experts. We selected it to ensure clinical and scientific relevance while avoiding selection bias that could arise from subjective article choice. The *Lancet* Seminar series thus offers authoritative, high-impact, and content-rich material, making it well-suited for expert-annotated sentiment training.

The abstracts were split into sentences (n=83), and were rated by four psychiatrists and four clinical psychologists. Ratings were provided on a visual analogue scale (0 = very negative and pessimistic, 100 = very positive and optimistic) using the open-source JavaScript application _magpie[3]. Each participant rated all sentences. The whole procedure took between 15 and 20 minutes. The sentences were presented in random order, and a mean score was calculated for each sentence. Sentences with a mean score of 50 or below were labeled as negative, while those with scores above 50 were labeled as positive. Inter-rater reliability was assessed using the intraclass correlation coefficient [ICC(2,k)] (Vallat, 2018), which indicated high agreement among raters, ICC = 0.879, 95% CI [0.83, 0.92], $p < .001$. Table 1 presents the descriptive statistics of the ratings. Since the number of sentences per class was imbalanced, we randomly selected 15 sentences from each class for the training dataset, resulting in a total of 60 sentences. The remaining 23 sentences were used for evaluation.

In addition, we constructed a test dataset of 200 synthetic sentences using GPT-4.0, with 50 sentences per label per condition, which we also used to evaluate model performance. We used the following prompt: *"Write 50 different sentences about [depression/schizophrenia] with [negative/positive] sentiment in the style of a scientific publication."* All synthetic sentences were verified by the authors for factual consistency and plausibility of the assigned labels. Furthermore, we applied a RoBERTa model fine-tuned for sentiment classification (Hartmann et al., 2023)[4]. This model correctly predicted the label of the positive sentences in 100% of cases and the negative labels in 90% of cases. The synthetic dataset has been made publicly available[5].

We fine-tuned two different sentence-transformer models—**sentence-transformers/all-mpnet-base-v2**[6] and **BAAI/bge-small-en-v1.5**[7] on an NVIDIA A100 GPU to predict to predict probabilities of negative and positive sentiment ranging between 0 and 1. For brevity, we refer to them as Model 1 and Model 2. We consider Model 1 particularly appropriate for our task, as it is partially trained on scientific text from the S2ORC (Lo et al., 2020) corpus and incorporates domain knowledge from the medical field, making it better suited to capture the nuanced sentiment expressed in abstracts on schizophrenia and depression. The contrastive learning during fine-tuning was performed using 1,860 unique sentence pairs. Our training data did not include sentences labeled as "neutral," but during fine-tuning, we used a lower learning rate (1e-6) and fewer epochs (n=10) to mitigate overconfidence in classification. Furthermore, to prevent biasing the models toward one of the two conditions, we included a preprocessing step in which the tokens "schizophrenia" and "depression" were masked with "[condition]" in the training dataset. This masking strategy ensured that sentiment predictions were not driven by the lexical identity of the condition itself but by the surrounding linguistic context. The goal was to

[3] https://magpie-ea.github.io/magpie-site/

[4] https://huggingface.co/siebert/sentiment-roberta-large-english

[5] https://github.com/ivan-nenchev/Sentiment_schizophrenia_depression

[6] https://huggingface.co/sentence-transformers/all-mpnet-base-v2

[7] https://huggingface.co/BAAI/bge-small-en-v1.5

Condition	Mean score (Std)	Positive sentences (n=)	Negative sentences (n=)
Schizophrenia	46.5 (16.6)	18	27
Depression	48.5 (20.2)	20	18
Total	47.4 (18.3)	38	45

Table 1: Ratings (mean and standard deviation) and labels for 83 sentences extracted from *Lancet* abstracts on schizophrenia and depression.

Evaluation dataset	A	P	R	F1
Model 1	.78	.79	.78	.78
Model 2	.82	.82	.82	.82
Test dataset	**A**	**P**	**R**	**F1**
Model 1	1.0	1.0	1.0	1.0
Model 2	.96	.96	.96	.96

Table 2: Accuracy (A), precision (P), recall (R), and F1 scores of the fine-tuned models on the evaluation and test datasets.

promote generalization and fairness, enabling the models to learn sentiment patterns applicable to both conditions without overfitting to either term. Table 2 demonstrates the models' performance on the evaluation dataset and the test dataset with synthetic sentences.

After fine-tuning, both models rated the token "schizophrenia" as slightly more negative than "depression". To mitigate this bias, we removed both tokens from all sentences prior to sentiment extraction. Sentiment analysis was then performed on the cleaned dataset, with abstracts segmented into sentences and each sentence evaluated using the fine-tuned models. For each abstract and model, we calculated four sentiment metrics based on predicted probabilities: mean and maximum negative sentiment across all sentences and mean and maximum positive sentiment. Furthermore, we evaluated the sentiment of the titles. Additionally, in an exploratory analysis, we mapped sentiment values to the relative positions of sentences within abstracts to identify patterns in sentiment distribution. This approach provided insights into where sentiment typically appears within the structure of scientific abstracts.

Examples of the sentiment evaluation of titles and sentences can be found in Table 6 the Appendix.

3.3 Statistical analysis

Given that sentiment analysis models produce probability estimates constrained to the interval [0,1], we implemented beta regression models using the statsmodels package (Seabold and Perktold, 2010), which are optimal for modeling proportional data with bounded continuous outcomes. We constructed separate models for negative and positive sentiment polarities, utilizing both mean and maximum sentiment scores aggregated at the abstract level as dependent variables. In addtion, we fitted beta regression models for the sentiment of the titles.

The predictors included year and clinical condition (depression versus schizophrenia). To address substantial temporal variation in sample sizes—ranging from approximately 160 to over 10,000 abstracts per year—we computed a weighted mean year using abstract counts as weights, then centered the year variable around this weighted mean. This approach ensured that years with larger, more reliable samples exerted proportionally greater influence on temporal trend estimation while maintaining model stability.

Beta regression was selected over linear regression for its capacity to handle the natural boundaries and distributional characteristics of probability data without requiring potentially problematic transformations. Additionally, we fitted separate precision models with identical covariates to account for heteroscedasticity in sentiment variance across time and clinical conditions. This analytical framework allowed us to quantify diachronic changes in sentiment expression while simultaneously examining differential patterns between psychiatric conditions.

4 Results

The following section presents the results of the statistical analysis of sentiment values generated by the two fine-tuned models. Across both models,

	Model		Mean	Std	Min	Max
Title	1	negative	.488	.106	.112	.853
		positive	.512	.106	.147	.888
	2	negative	.505	.062	.265	.729
		positive	.495	.062	.271	.735
Sentence	1	negative	.45	.083	.099	.845
		positive	.55	.083	.155	.901
	2	negative	.486	.047	.25	.727
		positive	.514	.047	.273	.75

Table 3: Descriptive results for negative and positive sentiment in titles and sentences extracted using Model 1 *(sentence-transformers/all-mpnet-base-v2)* and Model 2 (*BAAI/bge-small-en-v1.5*).

		Centralized years				Schizophrenia			
Sentiment	**Model**	β	SE	z	p	β	SE	z	p
mean	1	-.004	.000	-60.41	<.001	.048	.002	25.6	<.001
negative	2	-.002	.000	-58.19	<.001	.018	.0017	19.73	<.001
maximum	1	.004	.000	42.53	<.001	.031	.003	15.5	<.001
negative	2	.002	.000	49.70	<.001	.01	.001	11.62	<.001
mean	1	.004	.000	60.41	<.000	-.04	.002	-25.62	<.001
positive	2	.002	.000	58.19	<.001	-.018	.001	-19.73	<.001
maximum	1	.008	.000	101.76	<.001	-.05	.002	-25.005	<.001
positive	2	.006	.000	124.21	<.001	-.0257	.001	-23.561	<.001

Table 4: Results from the beta regressions for the sentiment estimations of sentences extracted using Model 1 *(sentence-transformers/all-mpnet-base-v2)* and Model 2 (*BAAI/bge-small-en-v1.5*).

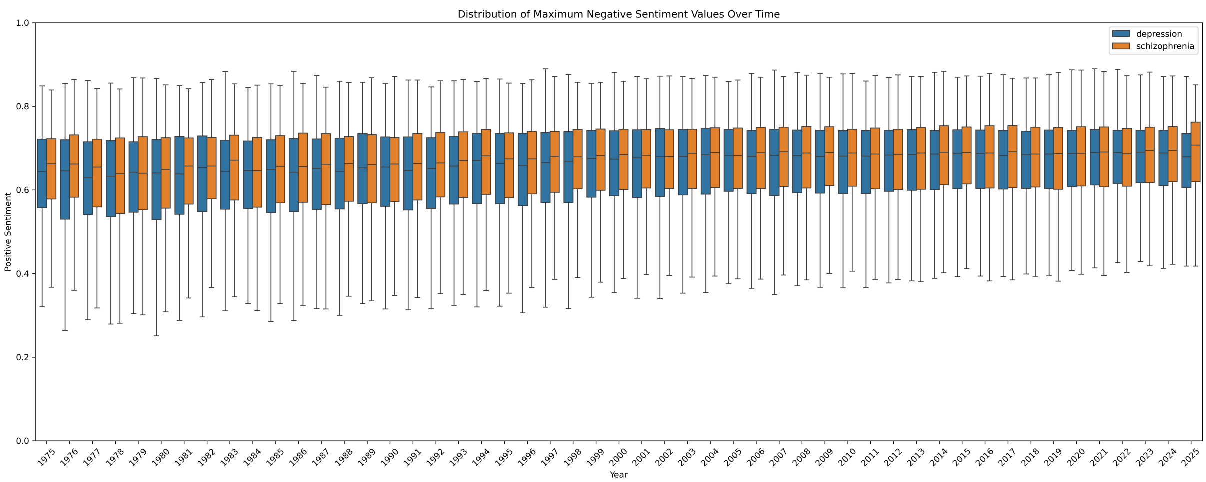

Figure 1: Increase in the maximum negative sentiment in abstracts on depression and schizophrenia.

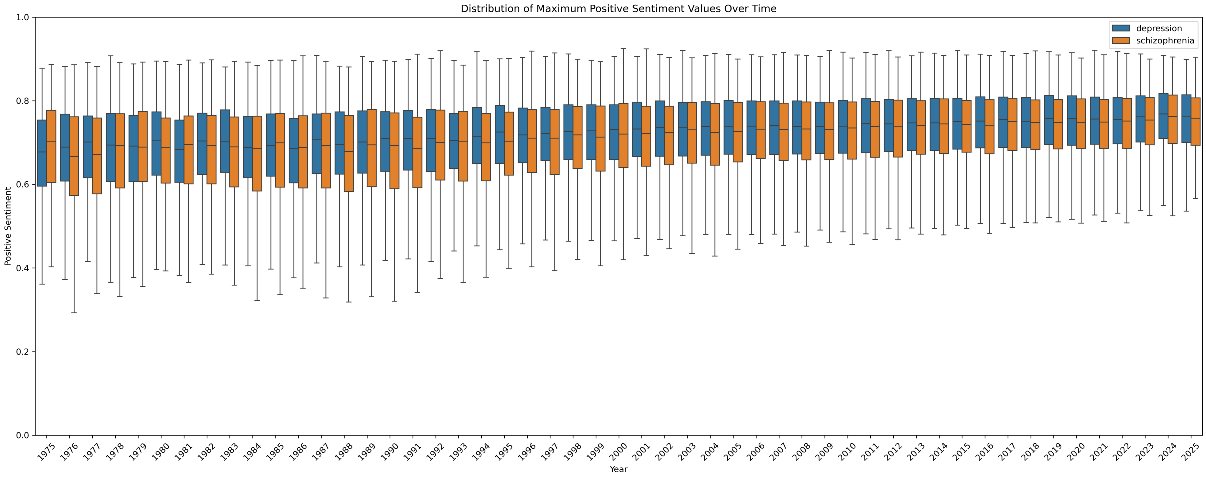

Figure 2: Increase in the maximum positive sentiment in abstracts on depression and schizophrenia.

beta regression analyses revealed that schizophrenia was associated with more negative sentiment than depression in both titles and abstracts. Over time, mean negative sentiment decreased, while the maximum values for both negative and positive sentiment increased. Descriptive results are presented in Table 3. Density plots of the probability distributions for positive and negative sentiment from Model 1 (sentence-transformers/all-mpnet-base-v2) are presented in Figure 5 in the Appendix.

The beta regression of sentiment values in titles from Model 1 revealed a decrease in negative sentiment (*β = -.0076, SE = .000, z = -104.90, p < .001*) and a corresponding increase in positive sentiment over time. Titles of papers on schizophrenia had a more negative sentiment compared to those on depression (*β = .07, SE = .002, z = 39.61, p < .001*). The results from Model 2 (BAAI/bge-small-en-v1.5) confirm the same pattern, with a significant effect of centralized year (*β = -.002, SE = .000, z = -52.94, p < 0.001*) and schizophrenia (*β = .032, SE = 0.000, z = 31.2, p < 0.001*). Figures 6 and 7 in the Appendix illustrate these findings.

The results from the beta regression models on the sentiment scores extracted on the sentence level of the abstracts are summarized in Table 4. Figures 1 and 2 illustrate the temporal changes in the maximum negative and positive sentiment.

A beta regression on the mean negative sentiment values from Model 1 (sentence-transformers/all-mpnet-base-v2) revealed a significant positive association between schizophrenia and negative sentiment (*β = .048, SE= .002, z= 25.6, p < 0.001*), with a slight reduction over time (*β = -.004, SE= .000, z= -60.41, p < .001*). We found a corresponding effect in the mean positive sentiment values from the model. The beta regression for mean negative sentiment scores obtained with model 2 (BAAI/bge-small-en-v1.5) revealed a significant negative association between the centralized year variable and negative sentiment (*β = -.002, SE = .000, z = -58.19, p < .001*), although the magnitude was smaller. Schizophrenia was associated with more negative sentiment (*β = .031, SE = .003, z = 19.73, p < .001*).

A beta regression on the maximum negative sentiment from Model 1 (sentence-transformers/all-mpnet-base-v2) showed a positive effect of centralized years (*β = .004, SE = .000, z= 42.53, p < .01*) and schizophrenia (*β = 0.031, SE= .003, z= 15.5, p < .001*). The odds ration for negative sentiment is 1.004, meaning that each year the probability for negative sentiment increases with 0.4%. Similarly, the beta regression for maximum positive sentiment indicated a significant positive effect of centralized years (*β = .008, SE = .000, z = 101.76, p < .001*), while schizophrenia was significantly negatively associated (*β = -0.05, SE = .002, z = -26.005, p < .001*). The log odds from Model 2 show the same pattern.

In addition, we examined the relationship between negative sentiment and relative position within abstracts, using average values for each year and position. Based on the visual exploration of the heatmap plot shown in Figure 3, we fitted an OLS model with a cubic term for position. The model explained 34% of the variance in sentiment scores *(R^2 = .34, F(3, 1016) = 174.5, p < .001)*. The linear term (*β = -.0390, p < .001*) indicated a general decrease in sentiment across positions,

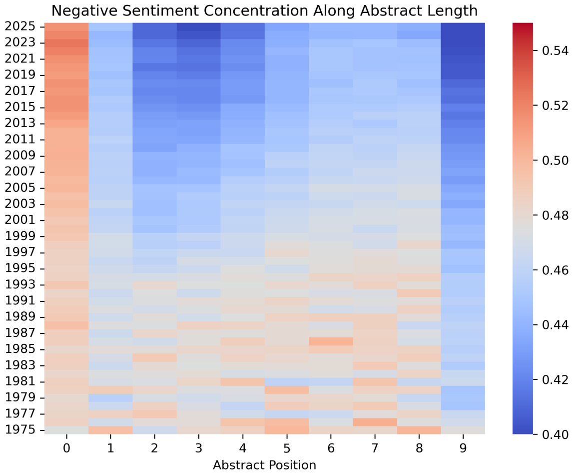

Figure 3: Heatmap of the relative position of positive and negative sentiment within abstracts.

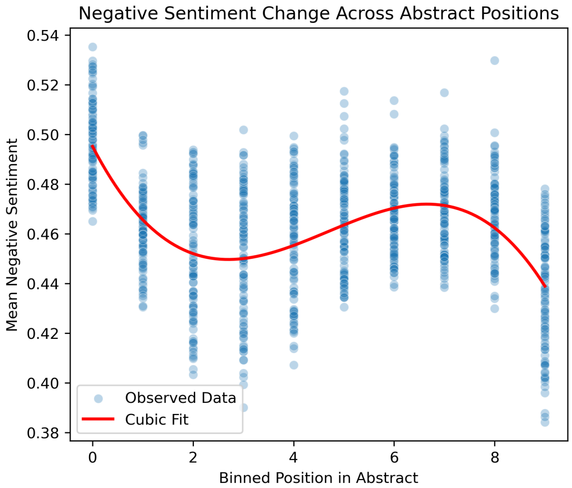

Figure 4: Scatter plot with a cubic regression curve.

while the quadratic (β = *.01, p < .001*) and cubic (β = -*.0007, p < .001*) terms suggest a more complex curvature.

5 Discussion

In this study, we analyzed sentiment related to schizophrenia and depression in a large corpus of PubMed abstracts published over the past 50 years. To achieve this, we adopted a domain-specific approach and fine-tuned two sentence-transformer models using the SetFit Python package. Our training dataset was derived from sentences in the *Lancet* Seminar series on schizophrenia and depression, which had been rated by psychiatrists and clinical psychologists. To minimize potential bias in both model fine-tuning and sentiment estimation, we excluded the tokens "schizophrenia" and "depression" from the linguistic data.

Our findings offer several key insights into sentiment trends. First, texts about schizophrenia tend to have a more negative tone. Research abstracts focused on schizophrenia exhibit significantly more negative sentiment across all regression models, even when "schizophrenia" and "depression" were excluded from both the training data and the linguistic data used for evaluation. At this stage, we cannot provide a definitive explanation for this pattern, but several possibilities emerge. One possibility is that the language associated with schizophrenia — for example in relation to symptoms such as delusions and hallucinations — is inherently more negative. Another explanation could be that certain research areas, such as treatment options and psychotherapy, could be more frequently studied in the context of depression, may be evaluated as less negative by the models. Additionally, linguistic style may reflect broader societal imbalances in perceptions of mental illness. There is evidence suggesting increasing acceptance of depression, while stigma toward schizophrenia continues to rise (Sittner et al., 2024; Schomerus et al., 2022).

Second, in terms of temporal analysis, we observed a significant increase of the mean positive sentiment for both titles and abstracts, suggesting that, on average, the tone of scientific writing has become more positive. Our results on the trend appear to align with previous work (Vinkers et al., 2015; Cao et al., 2021; Wen and Lei, 2022; Edlinger et al., 2023; Liu and Zhu, 2025; Hartmann et al., 2023) despite the different methodological approaches.

However, an in-depth analysis of the sentences revealed an interesting trend in extreme sentiment values: both maximum negative and positive sentiment have increased over time. The language has become more polarized, potentially reflecting a broader shift toward more passionate and assertive expressions in scientific discourse. This shift may be driven by the growing emphasis on research questions and the need to highlight the significance of findings. Additionally, our exploratory analysis showed that negative sentiment is often concentrated at the beginning of an abstract, while the most positive sentences tend to appear at the end. This pattern aligns with the findings of Edlinger et al. (2023), who observed that positive words are more frequent toward the end of abstracts. It also reflects the conventional structure of scientific abstracts, where researchers typically introduce a problem or knowledge gap at the outset and conclude with positive statements about solutions, contributions, or future research directions. This structural tendency mirrors the natural progression of scientific inquiry, from identifying a problem to presenting novel solutions.

Our findings are particularly relevant in the context of psychological constructs such as negativity bias and negativity dominance (Rozin and Royzman, 2001), which describe how negative information tends to carry more weight than positive information. Additionally, they align with the work of Bellucci (2023), who demonstrated that the sequential order of presenting information influences recall—specifically, that negative statements introduced first are more likely to be remembered. This raises the possibility that the sentences with negative sentiment within abstracts may overshadow their overall positive tone. Further research is needed to investigate how this shift in sentiment may impact impression formation, particularly among medical professionals, researchers, and the general public. A crucial question is whether this bias in sentiment could shape how scientific findings are interpreted, potentially influencing attitudes, clinical decision-making, and research priorities.

Ethical and societal implications

NLP provides powerful methodologies for analyzing vast amounts of linguistic data, allowing researchers to focus on specific aspects of text with precision. By uncovering patterns that may not be immediately apparent to human readers—due to the complexity and speed limitations of human cognition—NLP offers valuable insights into language use. In this study, we apply NLP for sentiment analysis of scientific abstracts on schizophrenia and depression, examining trends in scientific writing over the past decades. While our findings align with prior research in showing an overall increase in positive sentiment, we also identify two concerning patterns. First, both the titles and sentences of abstracts on schizophrenia exhibit more negative sentiment compared to those on depression. Second, we observe an increase in polarized language within abstracts, which may shape readers' impressions and inadvertently reinforce stereotypes. This linguistic trend underscores the need for greater awareness in scientific writing, encouraging researchers to critically reflect on their choice of language and avoid excessively negative framing. These patterns may also have implications for people affected by mental illness. If negative sentiment in scientific discourse contributes to the broader cultural narrative, it could influence public perceptions and potentially exacerbate self-stigmatization. Future research should examine whether such language trends affect how patients view themselves and their condition.

Limitations

This study has several limitations that should be considered when interpreting the results. First, the dataset used for fine-tuning was relatively small, which could have affected model performance. However, we addressed this issue by using SetFit, a framework specifically designed to achieve robust results with limited labeled data. Second, our fine-tuning process relied on only two sentiment labels ("positive" and "negative"), without a dedicated "neutral" category. This could have led to an artificial polarization of sentiment predictions. To mitigate this, we carefully adjusted hyperparameters during fine-tuning to prevent the model from becoming overly confident in assigning sentences to either category. Additionally, density plots of the models' predictions showed that most sentences were distributed near the center of the sentiment scale, suggesting that the models captured a more nuanced sentiment distribution despite the absence of an explicit "neutral" label. Thirdly, despite removing the words "schizophrenia" and "depression" from the training and evaluation data, other linguistic features associated with these topics may still introduce biases. The models may have learned to associate certain medical terms or research topics with sentiment in unintended ways. Lastly, changes in sentiment trends over time may be influenced by shifts in scientific norms and publication practices. Without controlling for these factors, it is difficult to determine whether the observed trends reflect actual changes in sentiment or broader shifts in academic discourse.

Ethical statements

The current study is part of a broader project on stigma and schizophrenia, which has received approval from the ethics board at Charité Universitätsmedizin Berlin. All data used in this analysis were obtained from publicly available PubMed abstracts, and no personally identifiable information was included.

Acknowledgements

This project was conducted as part of the study *Stigmatization and theory of mind in psychiatric professionals: Literature and visual arts as media for anti-stigma interventions in psychiatry (LIBIAS),* funded by the German Federal Ministry of Health. I.N. is a participant in the BIH Charité Digital Clinician Scientist Program funded by the Charité – Universitätsmedizin Berlin and the Berlin Institute of Health at Charité (BIH). The authors acknowledge the Scientific Computing of the IT Division at the Charité - Universitätsmedizin Berlin for providing computational resources that have contributed to the research results reported in this paper. The authors thank Dr. Daniel Schulze for valuable advice on the statistical analysis.

References

M. Alvarez-Jimenez, J. F. Gleeson, L. P. Henry, S. M. Harrigan, M. G. Harris, G. P. Amminger, E. Killackey, A. R. Yung, H. Herrman, H. J. Jackson, and P. D. McGorry. 2011. Prediction of a single psychotic episode: a 7.5-year, prospective study in first-episode psychosis. *Schizophrenia Research*, 125(2-3):236–246.

Naomi Baes, Henry Speagle, and Nick Haslam. 2022. Has Psychology Become More Positive? Trends in Language Use in Article Abstracts. *Frontiers in Psychology*, 13:870549.

Roy F. Baumeister, Ellen Bratslavsky, Catrin Finkenauer, and Kathleen D. Vohs. 2001. Bad is Stronger than Good. *Review of General Psychology*, 5(4):323–370. Publisher: SAGE Publications Inc.

Gabriele Bellucci. 2023. The organizational principles of impression formation. *Cognition*, 239:105550.

Xiaoke Cao, Lei Lei, and Ju Wen. 2021. Promoting science with linguistic devices: A large-scale study of positive and negative words in academic writing. *Learned Publishing*, 34(2):82–88.

Moritz Edlinger, Finn Buchrieser, and Guilherme Wood. 2023. Presence and consequences of positive words in scientific abstracts. *Scientometrics*, 128(12):6633–6657.

Su-Chen Fang, Cheng-Yi Huang, and Yu-Hsuan Joni Shao. 2022. Long-term Outcomes of Early Use of Long-Acting Injectable Antipsychotics in Schizophrenia. *The Journal of Clinical Psychiatry*, 83(4):21r14153.

Jochen Hartmann, Mark Heitmann, Christian Siebert, and Christina Schamp. 2023. More than a Feeling: Accuracy and Application of Sentiment Analysis. *International Journal of Research in Marketing*, 40(1):75–87.

Oliver D. Howes, Thomas Whitehurst, Ekaterina Shatalina, Leigh Townsend, Ellis Chika Onwordi, Tsz Lun Allenis Mak, Atheeshaan Arumuham, Oisín O'Brien, Maria Lobo, Luke Vano, Uzma Zahid, Emma Butler, and Martin Osugo. 2021. The clinical significance of duration of untreated psychosis: an umbrella review and random-effects meta-analysis. *World Psychiatry*, 20(1):75–95.

C. Hutto and Eric Gilbert. 2014. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1):216–225. Number: 1.

Xueying Liu and Haoran Zhu. 2025. The diachronic change in linguistic positivity in the academic book reviewing of language studies: a text-mining analysis. *Scientometrics*, 130(1):133–157.

Kyle Lo, Lucy Lu Wang, Mark Neumann, Rodney Kinney, and Daniel Weld. 2020. S2ORC: The Semantic Scholar Open Research Corpus. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4969–4983, Online. Association for Computational Linguistics.

Paul H Lysaker, Shirley M. Glynn, Sandra M. Wilkniss, and Steven M. Silverstein. 2010. Psychotherapy and recovery from schizophrenia: A review of potential applications and need for future study. *Psychological services*, 7(2):75–91.

Gin S. Malhi and J. John Mann. 2018. Depression. *Lancet (London, England)*, 392(10161):2299–2312.

Ida-Marie Molstrom, Julie Nordgaard, Annick Urfer-Parnas, Rasmus Handest, Jonas Berge, and Mads Gram Henriksen. 2022. The prognosis of schizophrenia: A systematic review and meta-analysis with meta-regression of 20-year follow-up studies. *Schizophrenia Research*, 250:152–163.

Ines Montani, Matthew Honnibal, Adriane Boyd, Sofie Van Landeghem, and Henning Peters. 2023. explosion/spaCy: v3.7.2: Fixes for APIs and requirements.

Joshua J. Myszewski, Emily Klossowski, Kristopher M. Schroeder, and Carrie A. Schroeder. 2022. Utilization of sentiment analysis to assess and compare negative finding reporting in veterinary and human literature. *Research in Veterinary Science*, 148:27–32.

Stefania Paolini, Meghann Gibbs, Brett Sales, Danielle Anderson, and Kylie McIntyre. 2024. Negativity bias in intergroup contact: Meta-analytical evidence that bad is stronger than good, especially when people have the opportunity and motivation to opt out of contact. *Psychological Bulletin*, 150(8):921–964.

Roy H. Perlis and David S. Jones. 2024. High-Impact Medical Journals Reflect Negative Sentiment Toward Psychiatry. *NEJM AI*, 1(1):AIcs2300066. Publisher: Massachusetts Medical Society.

Bernice A. Pescosolido, Andrew Halpern-Manners, Liying Luo, and Brea Perry. 2021. Trends in Public Stigma of Mental Illness in the US, 1996-2018. *JAMA Network Open*, 4(12):e2140202.

Lebogang Phahladira, Hilmar K. Luckhoff, Laila Asmal, Sanja Kilian, Frederika Scheffler, Stefan du Plessis, Bonginkosi Chiliza, and Robin Emsley. 2020. Early recovery in the first 24 months of treatment in first-episode schizophrenia-spectrum disorders. *npj Schizophrenia*, 6(1):1–8. Publisher: Nature Publishing Group.

Paul Rozin and Edward B. Royzman. 2001. Negativity Bias, Negativity Dominance, and Contagion. *Personality and Social Psychology Review*, 5(4):296–320. Publisher: SAGE Publications Inc.

Georg Schomerus, Stephanie Schindler, Christian Sander, Eva Baumann, and Matthias C. Angermeyer. 2022. Changes in mental illness stigma over 30 years – Improvement, persistence, or deterioration? *European Psychiatry*, 65(1):e78.

S. K. Schultz and N. C. Andreasen. 1999. Schizophrenia. *Lancet (London, England)*, 353(9162):1425–1430.

Skipper Seabold and Josef Perktold. 2010. Statsmodels: Econometric and Statistical Modeling with Python. *Proceedings of the 9th Python in Science Conference*, pages 92–96. Conference Name: Proceedings of the 9th Python in Science Conference.

M. Sittner, T. Rechenberg, S. Speerforck, M. C. Angermeyer, and G. Schomerus. 2024. 'Broken souls' vs. 'mad ax man' – changes in the portrayal of depression and schizophrenia in the German media over 10 years. *Epidemiology and Psychiatric Sciences*, 33:e37.

Yla R. Tausczik and James W. Pennebaker. 2010. The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology*, 29(1):24–54. Publisher: SAGE Publications Inc.

Lewis Tunstall, Nils Reimers, Unso Eun Seo Jo, Luke Bates, Daniel Korat, Moshe Wasserblat, and Oren Pereg. 2022. Efficient Few-Shot Learning Without Prompts. *arXiv preprint*. ArXiv:2209.11055 [cs].

Raphael Vallat. 2018. Pingouin: statistics in Python. *Journal of Open Source Software*, 3(31):1026.

Emilie Veerapa, Pierre Grandgenevre, Mohamed El Fayoumi, Benjamin Vinnac, Océanne Haelewyn, Sébastien Szaffarczyk, Guillaume Vaiva, and Fabien D'Hondt. 2020. Attentional bias towards negative stimuli in healthy individuals and the effects of trait anxiety. *Scientific Reports*, 10(1):11826. Publisher: Nature Publishing Group.

Christiaan H. Vinkers, Joeri K. Tijdink, and Willem M. Otte. 2015. Use of positive and negative words in scientific PubMed abstracts between 1974 and 2014: retrospective analysis. *BMJ*, 351:h6467. Publisher: British Medical Journal Publishing Group Section: Research.

G Nduku Wambua, Sanja Kilian, Vuyokazi Ntlantsana, and Bonginkosi Chiliza. 2020. The association between resilience and psychosocial functioning in schizophrenia: A systematic review and meta-analysis. *Psychiatry Research*, 293:113374.

Mayur Wankhade, Annavarapu Chandra Sekhara Rao, and Chaitanya Kulkarni. 2022. A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55(7):5731–5780.

Ju Wen and Lei Lei. 2022. Linguistic positivity bias in academic writing: A large-scale diachronic study in life sciences across 50 years. *Applied Linguistics*, 43(2):340–364.

Samantha E. Williams, Jaclyn H. Ford, and Elizabeth A. Kensinger. 2022. The power of negative and positive episodic memories. *Cognitive, Affective & Behavioral Neuroscience*, 22(5):869–903.

World Health Organization. 2016. *International Statistical Classification of Diseases and Related Health Problems, 10th Revision*, fifth edition. World Health Organization.

Appendix

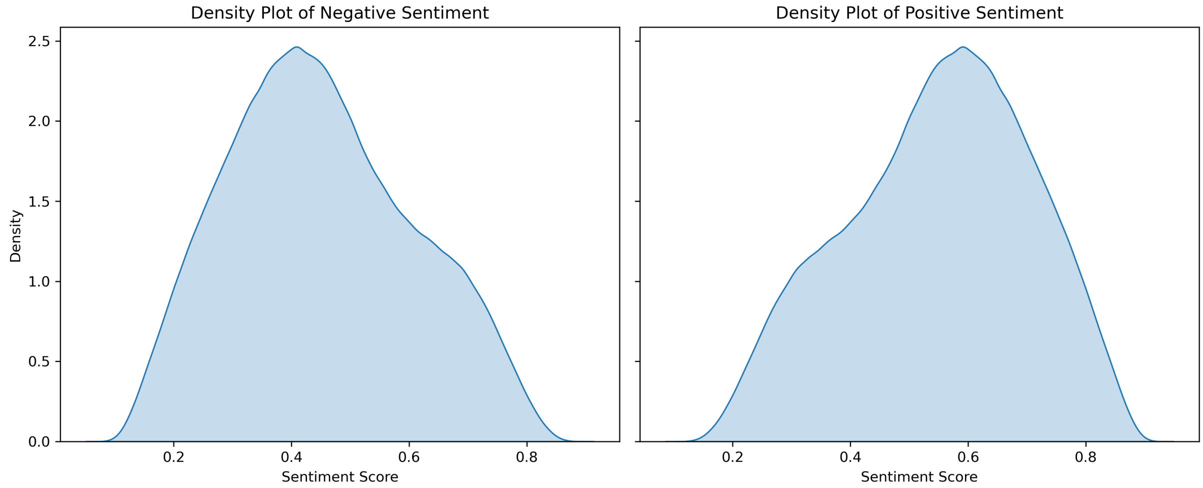

Figure 5: Density distributions of positive and negative sentiment (Model 1).

Figure 6: Negative sentiment in titles (Model 1).

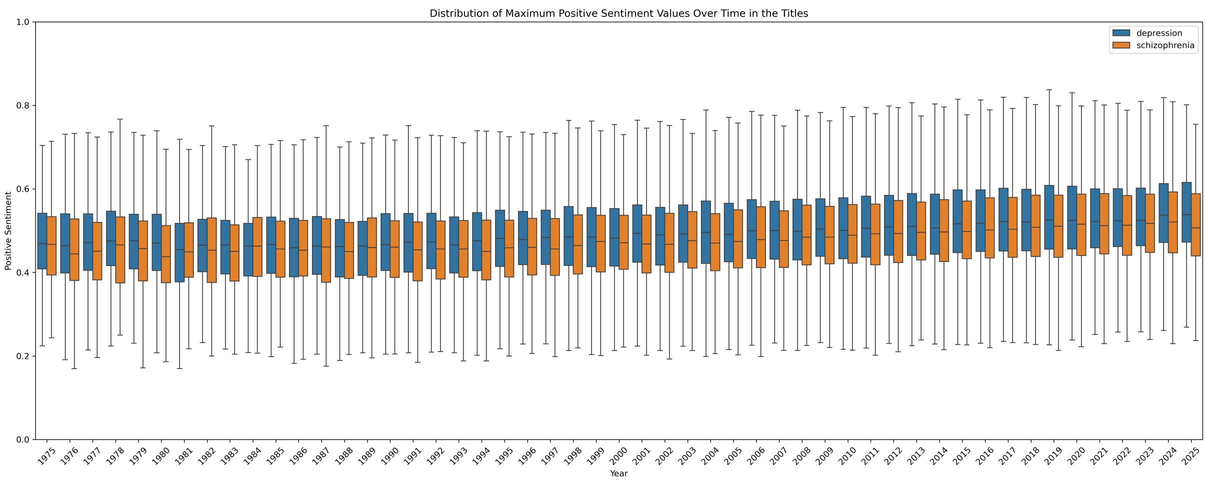

Figure 7: Positive sentiment in titles (Model 1).

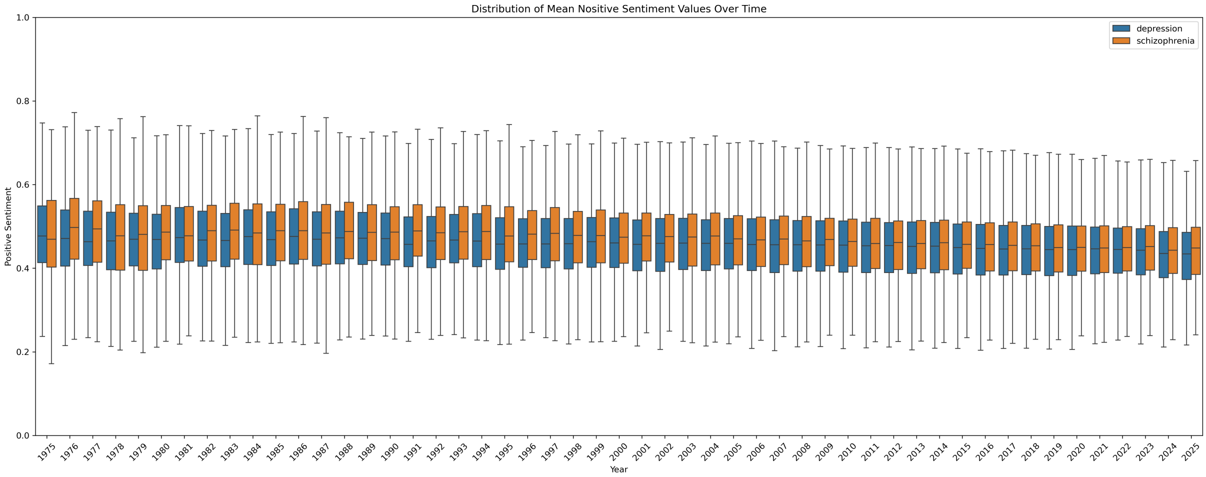

Figure 8: Mean negative sentiment per abstract (Model 1).

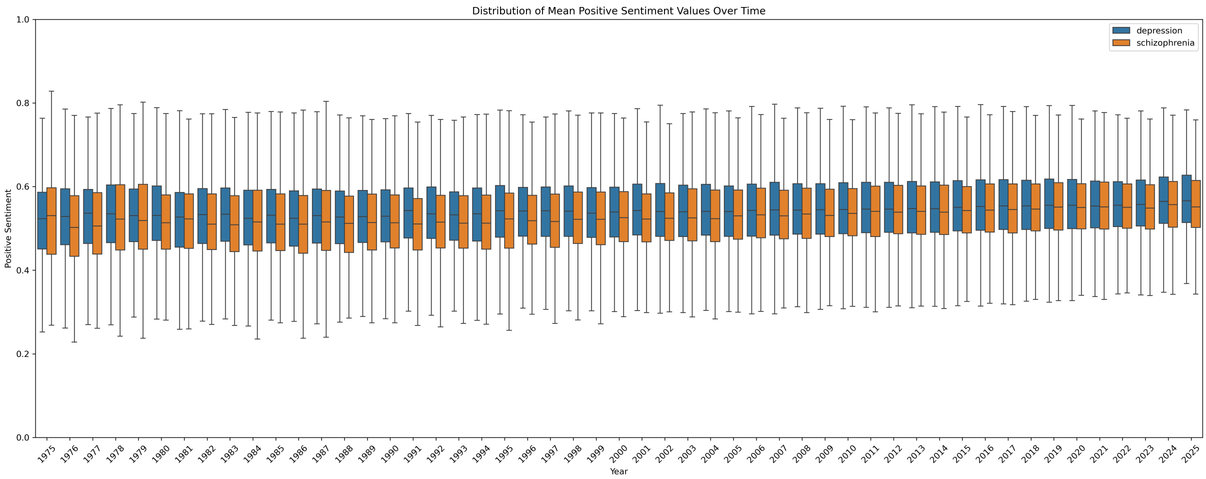

Figure 9: Mean positive sentiment per abstract (Model 1).

	Depression			Schizophrenia		
Year	**Abstracts n=177145**	**Tokens (mean)**	**Sentences (mean)**	**Abstracts n=83412**	**Tokens (mean)**	**Sentences (mean)**
1975	273	121.17	5.39	187	117.68	5.2
1976	529	123.63	5.53	460	122.9	5.42
1977	524	123.35	5.69	462	118.23	5.19
1978	496	112.38	5.15	371	113.3	5.06
1979	529	117.81	5.21	469	114	5.07
1980	532	119.83	5.51	446	115	5.23
1981	569	125.52	5.53	566	120.01	5.25
1982	515	135.21	6.16	534	123.59	5.57
1983	569	136.2	6.22	579	120.92	5.4
1984	645	129.86	5.69	634	121.02	5.34
1985	721	135.88	5.9	689	121.31	5.39
1986	732	133.53	5.78	665	123.26	5.41
1987	773	135	5.92	620	121.99	5.31
1988	795	137.85	6.05	587	123.82	5.42
1989	1020	143.96	6.41	797	126.79	5.72
1990	1002	142.28	6.24	775	128.68	5.7
1991	1163	149.03	6.49	961	132.29	5.86
1992	1297	149.82	6.71	1120	138.61	6.18
1993	1126	156.01	6.96	914	145.35	6.41
1994	1287	162.36	7.24	1077	151.22	6.74
1995	1403	162.48	7.24	1217	154.75	6.83
1996	1246	175.33	7.72	1057	159.17	7.11
1997	1880	174.59	7.78	1643	164.51	7.34
1998	1658	182.74	8.13	1338	171.49	7.62
1999	1944	184.85	8.38	1551	176.64	7.83
2000	1902	186	8.47	1404	186.06	8.48
2001	2248	189.45	8.64	1638	185.33	8.33
2002	2318	192.67	8.82	1514	184.98	8.42
2003	2603	192.16	8.74	1699	188.98	8.52
2004	3208	194.2	8.93	2251	190.11	8.56
2005	3393	195.77	8.96	2188	192.97	8.67
2006	3983	195.83	8.95	2382	192.72	8.63
2007	4348	195.45	8.97	2678	195.38	8.76
2008	4338	199.5	9.14	2527	195.65	8.79
2009	4913	200.79	9.16	2582	203.04	9.08
2010	4926	201.83	9.22	2425	202.11	9.01
2011	5938	204.3	9.39	2870	199.05	8.88
2012	6380	208.99	9.57	2948	205.49	9.19
2013	5928	215.16	9.9	2838	209.05	9.3
2014	7138	214.75	9.85	3037	211.01	9.49
2015	7285	221.31	10.16	2931	212.58	9.53
2016	7861	221.45	10.17	3169	212.29	9.48
2017	7629	224.65	10.38	2568	218.9	9.84
2018	8092	223.74	10.32	2870	213.57	9.66
2019	8303	228.15	10.57	2768	218.03	9.76
2020	8832	232.69	10.82	2642	221.14	9.87
2021	12792	231.71	10.84	3937	225.36	10.17
2022	9802	235.64	10.95	2979	228.57	10.24
2023	8114	238.87	11.11	2294	229.81	10.43
2024	10751	244.86	11.42	2387	236.81	10.81
2025	892	235.22	10.98	167	236.99	10.93

Table 5: Descriptive results of the PubMed corpus.

Condition	Negative Score	Year	Example
depression	.89	1997	Uncontrolled observations indicate that it could be associated with a remarkable deterioration in the course of the disease.
	.89	2021	It is characterized by a high recurrence rate, disability, and numerous and mostly unclear pathogenic mechanisms.
	.87	2007	If persistent, the condition can lead to significant disability.
	.49	1977	Reference is also made to the existence of various disease states where abnormalities of biogenic amines exist in the absence of affective disorders.
	.49	2025	Depressive symptoms were assessed using the Center for Epidemiological Studies [condition] Scale (CESD-10).
	.49	2015	A regular screening of such patients is thus essential for prognosis.
	.08	2010	Recent studies have shown the therapeutic value of the behavioral activation component of such interventions.
	.08	1998	Research into the efficacy of psychotherapy has often reported equivalence in treatment outcome when comparing different therapies.
	.09	2002	There is encouraging early evidence from multi-centre randomized controlled trials.
schizophrenia	.88	2014	In particular, compliance problems constitute a poor prognostic factor for this disorder due to increasing risk of relapse and hospitalization.
	.86	1989	The most frequent causes are patient's omission to take prescribed drugs, environmental conflicts and alteration in the familial situation.
	.86	1979	Habitual noncompliers have little investment in staying well, and cannot be expected to bear even mild drug side-effects.
	.49	1989	The battery consists of 26 items.
	.49	2022	The trends of digit span tests, correct number of consonants and inconsonant were increasing.
	.49	2001	When sentences moderately biased subordinate meanings (e.g., the animal enclosure meaning of pen), patients showed priming of dominant targets (e.g., paper) and subordinate targets (e.g., pig).
	.09	2000	A rich formulary of psychosocial interventions with demonstrated efficacy is now available.
	.08	1991	Some encouraging studies on efficacy are already available.
	.08	1996	Outcomes research on treatments for [condition] has identified a number of efficacious interventions.

Table 6: Examples of negative sentiment scores from Model 1.