Datenbank-Spektrum
https://doi.org/10.1007/s13222-026-00528-w

SCHWERPUNKTBEITRAG

Narratives and Linguistic Features of Drivel

Philipp Heinrich[1] · Andreas Blombach[1] · Stephanie Evert[1] · Fabian Schäfer[2]

Received: 11 October 2025 / Accepted: 16 January 2026

Abstract
Our contribution is concerned with the increasing amount of conspiracy theories and other forms of mis- or disinformation spreading on social media. We address this challenge to democratic opinion formation with a two-pronged approach. On the one hand, we identify specific conspiracy narratives in a social media corpus; on the other hand, we look at general linguistic dimensions that contribute to the overall drivel-like quality of such texts regardless of the narratives involved. For the present contribution, six distinct dimensions, along with an overall measure of drivel-like quality, were assessed using a five-point Likert scale. A sample of approximately 2000 texts drawn from German Telegram was manually annotated. We present the calculation of inter-annotator agreement to evaluate annotation consistency, conduct correlation analyses to examine the relationships between the individual dimensions, and fit a linear model to predict the overall drivel-like quality of the texts based on the individual dimensions. In addition, we train ordinal regression models to predict the values of each dimension from bag-of-*n*-grams representations. Finally, an analysis of feature weights identifies which *n*-grams serve as the most reliable indicators of each dimension.

Keywords Drivel · Conspiracy theories · Social media · CMC · Annotation · Prediction · Ordinal regression

1 Introduction & Related Work

Social media platforms are increasingly populated with conspiracy-related and conspiracy-adjacent content, fake news, disinformation, etc. Such content may originate from genuine user contributions, but it can also include state-sponsored propaganda, for instance from countries such as Russia, Israel, or China. Its proliferation poses a serious threat to democratic opinion formation [1]. Detecting such content is therefore important; the aim is not necessarily to censor users, but rather to provide clarification and ensure the dissemination of accurate information.

In prior publications [2, 3], we use the term "drivel"[1] for posts of this kind. Although "drivel" is a subjective label and lacks a precise definition in the existing literature, we employ it as a practical term for a recurring type of nonsensical, pseudo-profound, or obscure language commonly found in conspiracy-related or conspiracy-adjacent content. We distinguish between two complementary aspects of drivel. The first concerns content in the form of *specific narratives*—such as conspiracy theories, homeopathic claims, or religious assertions with real-world implications. The second concerns *stylistic linguistic characteristics*—including vague formulations, grandiose but imprecise statements, and language that presents speculation as certainty. Both aspects contribute to making drivel an intuitive and easily recognised category for most language users.

Identifying specific narratives is comparatively straightforward and is crucial for tasks like fact-checking and countering disinformation. Consequently, prior research has focused both on binary classification tasks, distin-

✉ Philipp Heinrich
philipp.heinrich@fau.de

Andreas Blombach
andreas.blombach@fau.de

Stephanie Evert
stephanie.evert@fau.de

Fabian Schäfer
fabian.schaefer@fau.de

[1] Chair of Computational Corpus Linguistics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Bismarckstr. 6, 91054 Erlangen, Germany

[2] Chair of Japanese Studies, Friedrich-Alexander-Universität Erlangen-Nürnberg, Artilleriestr. 70, 91052 Erlangen, Germany

[1] As a translation of German *Geschwurbel*.

Published online: 02 February 2026

guishing content that contains conspiracy theories or disinformation from content that does not, and on detecting particular narrative types, for example *COVID-19*-related conspiracies [2]. SemEval 2025 Task 10 [4] addresses the detection of narratives related to the *Russo-Ukrainian war* and *climate change*. Subtask 2 of this competition introduces a hierarchical classification problem in which texts must be assigned to narratives and sub-narratives. Many of these narratives directly reflect conspiratorial discourse, such as framing the West as the aggressor in the Russo-Ukrainian war or claiming that scientists disagree about the existence of anthropogenic climate change. Traditional machine-learning methods often struggle with such nuanced distinctions but fine-tuned large language models achieve strong performance [3]. The best-performing system [5] applies a hierarchical three-step prompting strategy: it first categorises an article into one of the two domains (*Russo-Ukrainian war* vs. *climate change*), then identifies the main narratives, and finally determines the appropriate sub-narratives. Section 3 gives a short overview of the narratives found in a corpus of conspiratorial posts from German Telegram (see Section 2).

The main focus of the present contribution, however, is on *linguistic features* that characterise drivel. Our aim is to investigate what exactly contributes to the drivel-like quality of texts that is intuitively recognised by language users. Analysis of inter-annotator agreement (see Section 5) in manual annotation tasks indicates that humans indeed have reliable intuitions about what makes texts "drivelly". In Section 6, we show that overall drivel-like quality of texts can be accurately predicted from specific *dimensions of drivel* found in the text, most notably its *distance from reality*, but also text-internal factors such as its *linguistic and argumentative peculiarities* or its *claim to absoluteness* (see Section 4). These dimensions help clarify how drivel operates at the level of language rather than content. In this sense, drivel can be understood as a *content-agnostic discursive feature*, analogous to how excessively complex language can serve as a marker of distinction rather than communication [6].

Defining the discursive features of drivel—that is, language that "spouts hot air", babbles without substance, or makes nonsensical claims—is a difficult endeavour. Intuitively, drivel shares important characteristics with what Frankfurt [7] terms *bullshit*: speech produced "without concern for the truth", which aims to convey an impression of meaning and sincerity. Linguistically, drivel often manifests as needlessly convoluted, vague, or pseudo-profound phrasing—language that pretends to be meaningful or intellectual while obscuring its actual content. As Frankfurt famously notes, "It is impossible for someone to lie unless he thinks he knows the truth. Producing bullshit requires no such conviction." [7] (55) This insight aligns closely with our conceptualisation of drivel as a stylistic strategy to some extend independent of actual truth.

We build upon a manual annotation of 1000 posts along the six dimensions introduced in [8], which rate texts according to criteria such as their distance from reality or their suggestiveness. We extend the annotation to approximately 2000 posts and report on inter-annotator agreement (Section 5), as well as the correlations between dimensions and their contributions to the overall drivel-like quality of the texts (Section 6). Finally, Section 7 addresses the prediction of individual dimension scores through ordinal regression using bag-of-*n*-gram representations, and presents an analysis of the linguistic features that serve as reliable indicators for each respective dimension.

2 Data

2.1 The *Schwurpus* Corpus

The data analysed in this contribution are sampled from the *Schwurpus* corpus, a collection of German conspiracy-theorist content sourced from Telegram [2]. Telegram is a minimally moderated messaging and microblogging platform. In Germany, many of the most prominent public channels are operated by conspiracy theorists[2]. This trend has been especially pronounced since 2020, as larger platforms such as YouTube, Twitter, and Facebook began removing, flagging, or demonetising channels spreading disinformation during the early stages of the COVID-19 pandemic [9, 10], although these efforts were considered somewhat insufficient [11]. More recently, Telegram has also become home to many cryptocurrency-related channels and other *finfluencer* accounts.

We downloaded the most prominent publicly available channels[3] in mid-2020 and again in mid-2022. To expand the corpus iteratively, we identified channels that were most frequently mentioned or linked to in the original corpus. The final *Schwurpus* corpus comprises over 200 channels, with follower counts ranging from a few thousand to more than 300,000, as well as more than 100 public group chats. The data set spans a period from January 2020 to July 2022, encompassing over 13 million posts and nearly 400 million tokens. The full corpus is publicly accessible[4].

[2] See e.g. https://telemetr.io/en/catalog/germany.

[3] We use the term "channel" here to refer both to Telegram channels, i.e. chatrooms where only administrators can broadcast, and to groups, i.e. chatrooms where every member can post.

[4] Available upon request at https://corpora.linguistik.uni-erlangen.de/cqpweb/schwurpus_v2.

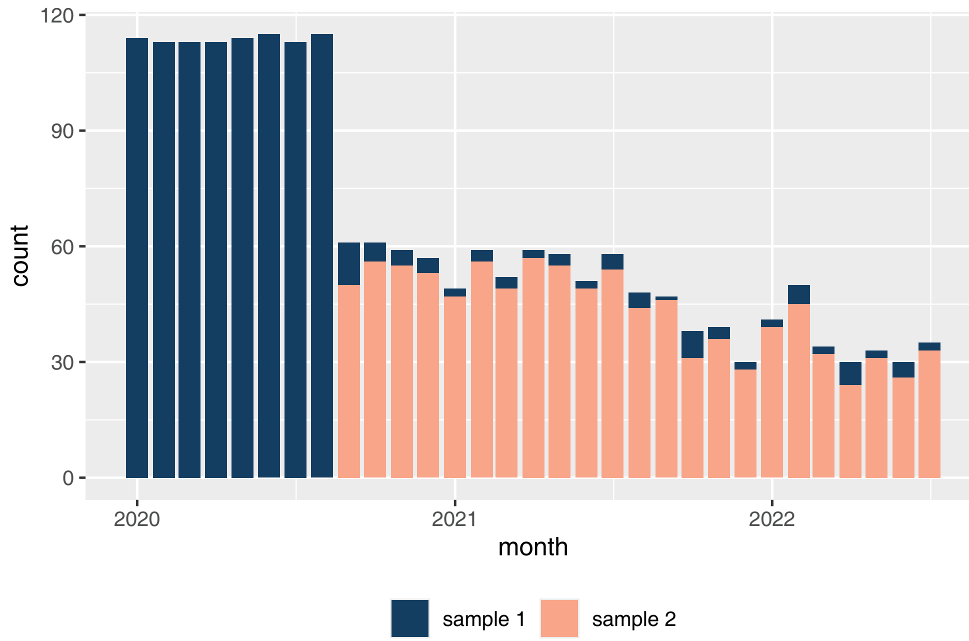

Fig. 1 Number of posts across time in both samples

2.2 Samples for Manual Annotation

In the present contribution, we extend the manual annotation presented in [8], which originally covered 1000 posts mostly from the first half of 2020 and only very few posts from the remainder of the period covered by *Schwurpus*. This initial set is referred to as *sample 1* in Fig. 1, which illustrates the distribution of posts in manually the annotated samples across time.

We augment this data set with an additional 1000 posts drawn uniformly from mid-2020 to mid-2022, designated as *sample 2* in Fig. 1. As in the first sample, only posts with 400 or more characters were included, and to ensure balanced representation and avoid bias towards highly active channels, the data were stratified by month and by channel frequency category.

The posts in both samples originate from a total of 238 different channels, with the post counts of the most prominent channels reported in Table 1. The numbers add up to 993 and 996 texts, respectively; although we did deduplicate the corpus before sampling—based on the algorithm presented in [12]—we had to manually filter out further duplicates during the annotation stage.

Table 1 Number of posts from the most prominent channels in each sample.

	Sample		
Channel	1	2	Total
Evahermanoffiziell	50	7	57
Qlobalchange	42	10	52
Uncutnewsschweiz	26	22	48
Kulturstudio	39	7	46
Alternativemedien	40	3	43
Oliverjanich	35	7	42
Wirsindvielmehr	31	8	39
Uncut_news	19	18	37
Wir_sind_viel_mehr	28	7	35
Coronavirushilfe	21	11	32
Expresszeitung	28	4	32
q_anonymous_kanal_deutschland	17	14	31
Epochtimesde	20	7	27
holistischegesundheitheilung	23	4	27
Klagemauerwand	20	6	26
Attilahildmann	16	5	21
Gwisnewski	9	11	20
Unzensiert	16	3	19
Efauf	14	4	18
Freiemedientv	10	8	18
Other	464	830	1294
Total	993	996	1989

3 Narratives

In [2], we developed a hierarchical narrative categorisation scheme based on prior research [13, 14], domain expertise, and close reading of corpus excerpts. The scheme comprises 18 narrative groups subdivided into 63 fine-grained narratives, each with descriptive sentences and examples, available online. Description sentences, derived from expert annotation guidelines, summarise each narrative concisely. The scheme includes COVID-19-specific narratives (e.g. claims that *COVID-19 is no more dangerous than the common flu* or that *the pandemic serves to implement the*

Great Reset) as well as pre-existing narratives such as *New World Order* and *sheeple*.[5]

Narratives surrounding COVID-19 comprise a broad range of theories suggesting that the pandemic was engineered or manipulated to advance hidden agendas, such as global surveillance through microchipping, profit-making by pharmaceutical elites, or facilitating the so-called Great Reset. Narrative groups include *pseudo-pandemic narratives*, which downplay the danger or even deny the existence of COVID-19, for example by equating it with seasonal influenza or suggesting that case numbers are exaggerated for political purposes. *Criticism of countermeasures* encompasses claims that pandemic restrictions are unlawful, authoritarian, or more harmful than the virus itself, such as assertions that mask mandates cause illness or that lockdowns were intended to control the population rather than protect public health. *Alternative treatments* refer to alleged cures or preventatives—for instance, ivermectin, hydroxychloroquine, or homeopathic remedies—that are said to be suppressed by governments and pharmaceutical companies. *Vaccine hazards* narratives portray COVID-19 vaccines as dangerous or part of an experimental mass-testing programme, with recurrent themes of infertility, DNA alteration, or deliberate population control.

These narratives are often interlinked with other narratives unrelated to COVID-19, including claims about chemtrails, mind control, or secret societies orchestrating world events. *QAnon*-related narratives allege that an elite cabal engages in child trafficking and that a covert struggle between good and evil is underway, with political figures such as Donald Trump portrayed as saviours. Further narrative clusters reflect ideological or social worldviews; *group-focused enmity* includes racist, xenophobic, misogynist, or antisemitic content as well as far-right myths like the *Great Replacement* or the *BRD GmbH* theory, which denies the legitimacy of the modern German state. The *sheeple* narrative depicts the general population as gullible and blindly obedient to authority. *Millenarianism* anticipates a coming moment of reckoning or enlightenment when the "truth" will be revealed and believers vindicated. *State as an enemy* narratives question the existence of democracy, alleging a "deep state" or the manipulation of elections. *Indoctrination* narratives claim that mainstream media and educational institutions are controlled by powerful elites to suppress dissenting opinions. Finally, *esotericism and pseudo-science* encompasses anti-scientific and mystical beliefs, including the denial of viruses, faith in energy healing, and reliance on astrology or spiritual awakening. Other low-prevalence narratives include themes such as *climate change denial*.

[5] Some fine-grained categories may be more appropriately considered building blocks of narratives rather than complete narratives.

In our previous work [2], we showed that these narratives can be detected automatically in corpora with fairly good accuracy via zero-shot learning based on our narrative descriptions, offering the prospect of scaling research on disinformation narratives to very large corpora.

4 Dimensions of Drivel

In the present study, we focus on features of drivel largely independent of actual content, aiming to characterise what makes a text "drivelly" without explicit reference to particular topics or narratives (such as COVID-19 conspiracy theories), opening up a broader research perspective on drivel beyond specific conspiratorial discourses. The following description of six dimensions of drivel follows [8].

Dimension 1: Distance from Reality This dimension assesses how far a text departs from accepted reality and scientific consensus. Low scores indicate fact-based or experience-based statements that require no assumptions, while high scores correspond to entirely fabricated, fantastic, or conspiratorial claims. Intermediate values capture reliance on questionable premises, unverifiable beliefs, or religious and spiritual assertions with real-world implications.

Dimension 2: Linguistic and Argumentative Peculiarities This dimension evaluates the clarity and logical coherence of argumentation. Texts receiving low scores are clear, logical, and coherent, or non-argumentative, such as typical social exchanges. High scores indicate incoherent, illogical, or meaningless argumentation, including logical gaps, grammatical or spelling errors, associative reasoning, informal fallacies, personal attacks, and clickbait features.

Dimension 3: Claim to Absoluteness and Handling of Sources This dimension measures the strength of assertion and treatment of sources. Low scores reflect a cautious, balanced tone with hedging and openness to alternative views. High scores indicate a dogmatic, self-assured tone, often with manipulative or selective use of evidence. The highest level corresponds to an absolutist stance with unquestioned beliefs and a missionary zeal.

Dimension 4: Suggestiveness This dimension examines the extent to which a text subtly leads readers to unjustified conclusions. Features include subtext, dogwhistling, rhetorical questions, framing, loaded language, and simplified dichotomies. Low scores denote an absence of suggestive features, whereas high scores reflect strong and deliberate use of manipulative cues.

Dimension 5: Oversimplification This dimension assesses the reduction of complexity in argumentation. Low scores indicate nuanced, multifaceted discussions that consider major contributing factors. High scores reflect extreme reductionism, presenting single causes for complex issues, ignoring counterarguments, and distorting understanding. Intermediate scores capture varying degrees of simplification.

Dimension 6: Emotionality This dimension evaluates the emotional tone and expressiveness of the author. Features include emotionally charged vocabulary, expressive emojis, dramatic punctuation, and full capitalisation. Low scores indicate a neutral, objective tone, whereas high scores denote highly emotional or agitated text in which affect dominates the content.

5 Annotation

For the present analysis we compare annotations for the 1000 additional posts from sample 2 (see Section 2) with the previous annotations of sample 1 [8]. New annotators were initially trained on a subset of sample 1 to familiarise themselves with the annotation guidelines established in [8]. Manual adjudication was conducted during regular meetings and in consultation with the corresponding authors of this paper—focusing on those posts and dimensions that exhibited notable disagreement, as measured by the standard deviation of annotations. This procedure resulted in a total of 255 adjudicated values (compared to 198 adjudicated values in sample 1). All remaining annotations were resolved automatically by taking the means across annotators. Note that besides the six dimensions described above, annotators of both samples were also instructed to intuitively annotate the overall drivel-like quality of posts.

5.1 Distribution of Scores

Figure 2 presents the distribution of scores assigned by the new annotators in sample 2 on the right-hand side, alongside the distribution of scores per annotator in sample 1 on the left-hand side. The figure illustrates that annotators employ different strategies when assigning scores. Annotator 5 in sample 2 e.g. often assigned lower values, a pattern also observed for annotator 2 in sample 1. Very high scores

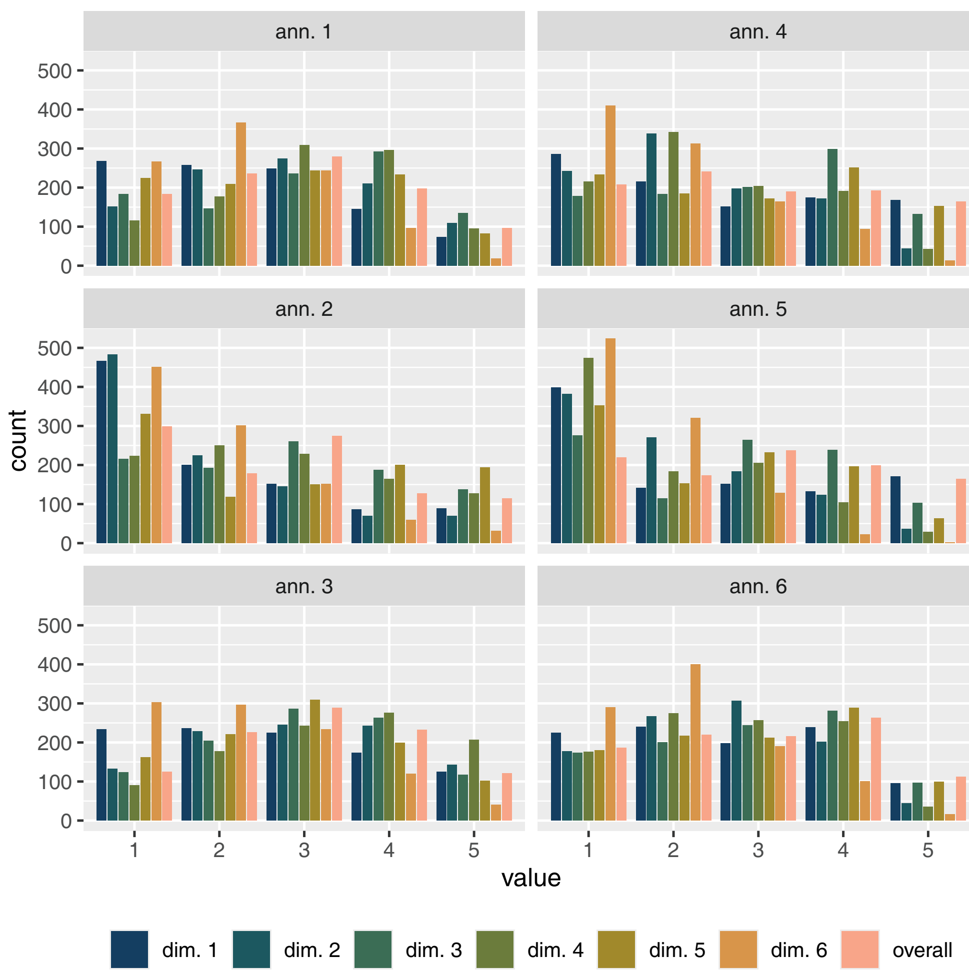

Fig. 2 Distribution of annotated scores by annotator and dimension. The left hand panels show annotators 1–3 on sample 1, the right hand panels annotatoers 4–6 on sample 2

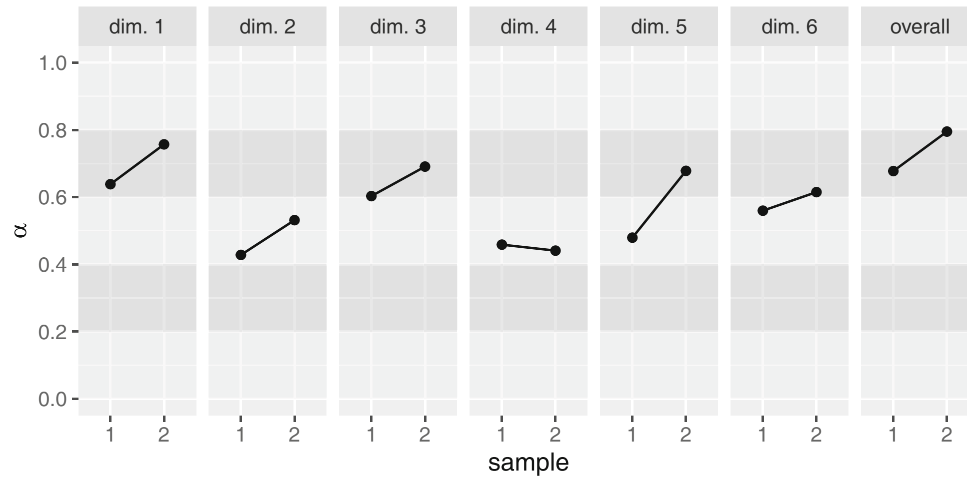

Fig. 3 Development of Krippendorff's α for three annotators and each dimension between samples. Grey bands represent the interpretation intervals

were rarely selected by any annotator. In an ordinal setting, achieving high inter-annotator agreement (IAA) can be easier when middle values are systematically chosen; however, annotators were explicitly instructed to make accurate judgements.

5.2 Inter-annotator-agreement

Following [8], we analyse IAA using Krippendorff's α for ordinal data [15]. In this setting, categories correspond to ranks, and α accounts for the order of categories by employing a distance function d_{ij} that reflects the distance between categories i and j. Here, we use the squared difference of ranks as the distance function. Krippendorff's α can be calculated for an arbitrary number of annotators. It ranges from -1.0, indicating perfectly discordant annotations, to $+1.0$, indicating perfect agreement, with negative values denoting agreement worse than chance. The interpretation of α values follows commonly accepted guidelines [15]: values between –1.0 and 0.0 indicate poor agreement; 0.0 to 0.2 slight agreement; 0.2 to 0.4 fair; 0.4 to 0.6 moderate; 0.6 to 0.8 substantial; and 0.8 to 1.0 near-perfect agreement.

Figure 3 compares IAA for all three annotators in sample 1 and sample 2, respectively. Overall, we report similar, yet slightly higher, agreement scores than in [8]. Annotators achieved substantial agreement for all dimensions except dimensions 2 (*linguistic and argumentative peculiarities*) and dimension 4 (*suggestiveness*), where they only exhibited moderate agreement. Dimension 4 is also the only dimension where agreement slightly decreased from sample 1 to sample 2. For the overall drivel-like quality, annotators reached near-perfect agreement in sample 2.

Pairwise IAA was especially high for annotators 4 and 6, with an average $\overline{\alpha} \approx 0.85$ across all dimensions (minimum 0.76 for dimension 2 and maximum 0.91 for overall drivel), while lower agreement was observed between annotators 5 and 6 ($\overline{\alpha} \approx 0.52$) and annotators 4 and 5 ($\overline{\alpha} \approx 0.57$). The lowest agreement scores were recorded for dimension 4, with $\alpha \approx 0.21$ between annotators 5 and 6, and $\alpha \approx 0.33$ between annotators 4 and 5.

6 Prediction of Overall Drivel-like Quality from Dimension Scores

To illustrate how individual dimensions contribute to the overall drivel-like quality of posts, we conduct a simple linear regression analysis.

6.1 Results

Table 2 presents results for two models. The first model includes all six dimensions, while the second model considers only those dimensions that can be predicted directly from the text itself, without requiring external world knowledge: *linguistic and argumentative peculiarities, claims to absoluteness*, and *emotionality*. The full model demonstrates that the overall drivel-like quality can be predicted with high accuracy from the individual dimensions ($R^2 = 0.93$), with *distance from reality* emerging as the single most important predictor, whereas *emotionality* contributes relatively little. The reduced model retains substantial explanatory power ($R^2 = 0.85$), with *claims to absoluteness* becoming the most influential factor, and *emotionality* not reaching significance.[6]

[6] Note that these models contain only seven and four parameters, respectively, in contrast to typical models used in NLP. The values of R^2 and adjusted R^2 are nearly identical, and overfitting is not a concern; we therefore refrain from reporting cross-validation results. In any case, the model's predictive value is limited, as it cannot be applied in practice without prior annotation of the dimensions.

Table 2 Full regression model (left) and model with reduced number of predictors (right). Three asterisks (***) indicate $p < 0.01$.

	Dependent variable:	
	Overall drivel-like quality	
	Model 1	Model 2
Dim. 1: distance from reality	0.514*** (0.012)	–
Dim. 2: ling. & arg. peculiarities	0.083*** (0.013)	0.282*** (0.018)
Dim. 3: claim to absoluteness	0.207*** (0.015)	0.755*** (0.015)
Dim. 4: suggestiveness	0.142*** (0.010)	–
Dim. 5: oversimplification	0.133*** (0.013)	–
Dim. 6: emotionality	0.027*** (0.010)	0.016 (0.015)
Constant	−0.097*** (0.023)	−0.121*** (0.032)
Observations	1991	1991
R^2	0.931	0.845
Adjusted R^2	0.930	0.845
Residual Std. Error	0.316 (df = 1984)	0.473 (df = 1987)
F Statistic	4441.315*** (df = 6; 1984)	3616.896*** (df = 3; 1987)

6.2 Correlation Analysis

In addition to regression analysis, it is informative to examine the correlations between individual dimensions and with the dependent variable. Figure 4 shows pairwise correlations for all dimensions. The strongest associations are observed between dimensions 1 (*distance from reality*) and 3 (*claim to absoluteness*), which are also highly correlated with the overall drivel-like quality and were identified as the main contributors in the regression models discussed above. These two dimensions also exhibit substantial correlation with dimension 5 (*oversimplification*). In contrast, the correlations of dimensions 6 (*emotionality*) and 4 (*suggestiveness*) with other dimensions and with the dependent variable are comparatively low.

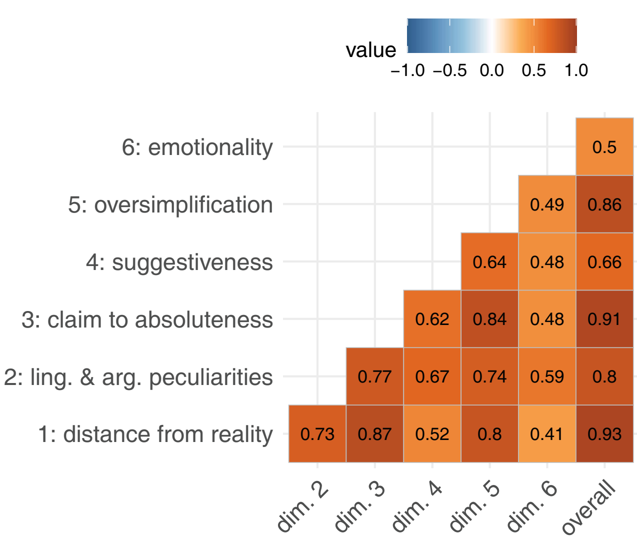

Fig. 4 Pearson correlation coefficients between dimensions and overall scores

7 Prediction of Each Dimension from Texts

The task of predicting dimension scores involves labels drawn from a finite, ordered set, here corresponding to ratings from 1 to 5. We employ classic ordinal regression using a bag of *n*-grams as features, although large language models might perform better in this setting. The motivation for using a simple linear machine learning algorithm is that it yields directly interpretable feature weights.

The bags of *n*-grams were created using the `scikit-learn` vectoriser [16]. Preprocessing steps comprised case folding and the exclusion of *n*-grams that appeared in more than 95% of documents. The feature space was restricted to a maximum of 100{,}000 dimensions. The vectoriser was configured to extract uni-, bi-, and trigrams, allowing the models to exploit both individual lexical cues and multi-word patterns. We use the standard tokeniser of the `scikit-learn` vectoriser, which yields only actual words and excludes emojis, punctuation marks, and other non-alphabetic characters.

Ordinal regression was implemented via the `mord` package [17]. In particular, we employed a classifier implementing the ordinal logistic model (Immediate-Threshold variant) [18], which accounts for the aspect that misclassifications further from the true label are more severe. For each dimension, we trained and evaluated a separate model in 5-fold cross-validation using the above features.

7.1 Evaluation

Predictions of the individual dimensions were evaluated using the Mean Squared Error (MSE), i.e. $\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$, where y_i and $\hat{y}_i$ denote the true and predicted scores, respectively, and N is the number of predictions.

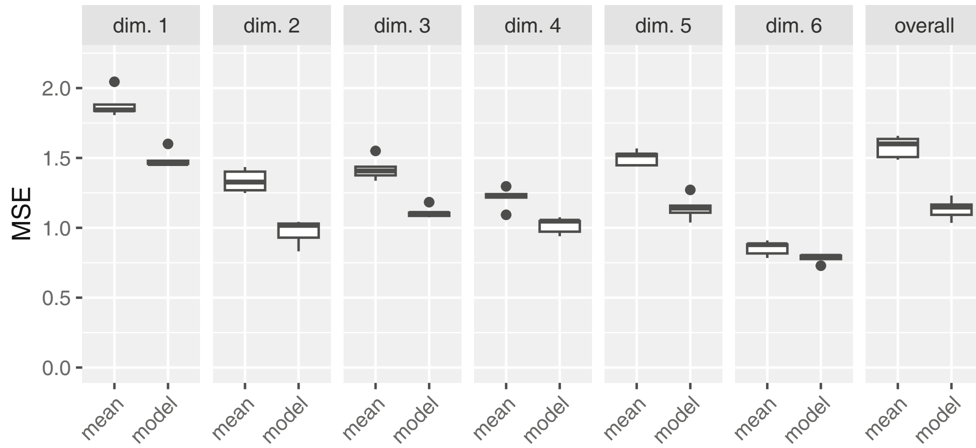

Fig. 5 Boxplots of the performance of our models for each dimension in 5-fold cross-validation, compared to a classifier based on the mean of scores in the training data

Figure 5 shows the performance of our models for each dimension in 5-fold cross-validation compared to a strong baseline, which is gained by using the mean of the dimension scores in the training data; our model outperforms it across all dimensions. The model shows particularly good performance on dimension 6; note however that this dimension is also the easiest to predict due to its score distribution, as reflected by baseline performance. Dimension 1 is the most challenging, both for the baseline and for the model's predictions.

These findings largely support our assumptions in Section 6.1, where we stipulated that dimensions 2, 3, and 6 can be accurately predicted from textual features without additional world knowledge. Moreover, the results indicate that dimension 4 (*suggestiveness*), which was not included in the reduced model, can also be accurately predicted from text.

7.2 Feature Weight Analysis

We conduct a feature weight analysis to identify which linguistic patterns contribute most to the correct prediction of each dimension. Our attention is on dimensions 2 (*linguistic and argumentative peculiarities*), 3 (*claim to absoluteness*), and 6 (*emotionality*), which we assumed to be most accurately predicted from bags of *n*-grams. We examine the features with the highest weights associated with high scores, averaging across all cross-validation folds. The most important features are unsurprisingly unigrams; however, these are more difficult to interpret compared with higher-order *n*-grams. For interpretability, we therefore concentrate on bigrams here, which are generally less context-sensitive than unigrams and can thus be interpreted more straightforwardly.

We consider only the top-50 bigrams for each of the three dimensions, which yields 108 distinct bigram types in total. Table 3 categorises these bigrams according to their membership in groups indicating overlap between dimensions. The most overlap occurs between dimensions 2 and 3, with less overlap between dimensions 2 and 6 and dimensions 3 and 6.

We further classify bigrams into three groups: functional, generic, and specific. Functional bigrams primarily serve a grammatical purpose, contain little or no semantic content, and typically include pronouns, articles, determiners, and auxiliary verbs. Their meaning depends on surrounding context and they generally express grammar rather than meaning. Generic bigrams refer to broad or ambiguous concepts, contain some semantic content but lack specificity, and do not identify a unique entity. Specific bigrams refer to unique, identifiable entities, such as people, organisations, places, events, URLs, or ideological constructs, and typically include proper nouns or named entities.

Our focus is on the interpretation of functional and generic bigrams. Content-specific bigrams in Table 3 are interesting in their own right and can often be linked to the narratives presented in Section 3. For example, bigrams associated with dimensions 2 and 3 reference well-known figures and events commonly associated with conspiracy theories (e.g. *bill gates, deep state*). In contrast, functional and generic bigrams do not indicate specific topics, but reflect stylistic or rhetorical choices, often employed for persuasion, repetition, or textitasis.

We acknowledge that the following interpretation constitutes an approach similar to that taken by corpus-assisted discourse studies (CADS) [19]. Given the crude preprocessing including removal of all non-alphabetic characters and excluding lemmatisation, it should however be considered a somewhat superficial exploration rather than a thorough linguistically informed analysis.

For dimension 2 (*linguistic and argumentative peculiarities*), bigrams indicate textitasis and certainty, such as *es gibt* and *das beste*. Temporal references appear in bigrams like *wenn es* and *plan ist*, while inclusion of the audience is reflected in bigrams such as *niemand möchte*. For dimension 3 (*claim to absoluteness*), bigrams expressing certainty

Table 3 Bigrams with highest feature weights for dimensions 2 (*linguistic and argumentative peculiarities*), 3 (*claim to absoluteness*), and 6 (*emotionality*).

Dimension	Bigrams
2 & 3 & 6	**Functional**: *das ist, es ist, es wird, ihr werdet, in die, und die, und nicht, und wenn*; **generic**: *die wahrheit*
2 & 3	**Functional**: *aus dem, das was, ist ein, nicht die, sie die, und das*; **generic**: *der plan, die menschen, die menschheit, report vom, zu machen*; **specific**: *22 report, bill gates, blogspot com, com 2020, deep state, der deep*
2 & 6	**Functional**: *das alles, ist der, nicht so*; **generic**: *ich glaube*
3 & 6	**Functional**: *nicht mehr, wenn der*; **generic**: *das system*
Only 2	**Functional**: *es aber, es gibt, es sich, euch nicht, ist ja, ist nicht, ja auch, mal auf, und so, und zu, wenn es, über den*; **generic**: *das beste, denke an, niemand möchte, plan ist*; **specific**: *die eu, https twitter, of the, twitter com*
Only 3	**Functional**: *gegen die, kann man, und der, wenn die, wenn ihr, werden diese, wir es*; **generic**: *der erde, der menschheit, der welt, die medien, die straße, verbrechen gegen*; **specific**: *change blogspot, die massenmedien, dirk dietrich, great reset, https qlobal, qlobal change, von dirk, übersetzung von*
Only 6	**Functional**: *an die, auch noch, auf und, das wird, du bist, es geht, euch an, für den, habe ich, hat es, ich hoffe, in einer, lasst euch, mal wieder, um die, und aus, und sich, was für, werden dann, wie viele, wir alle, wir mal, wird in, zu lassen*; **generic**: *der geschichte, die täter, die welt, diese menschen, ihre kinder, mehr informationen, menschen in, teilen teilen*; **specific**: *2020 01, me stuttgartgrundgesetzdemos*

include *ist ja*, while inevitability is signalled by *ihr werdet* and *werden diese*. Generalisations are evident in bigrams such as *die menschheit* and *die medien*. These bigrams indicate phrases that lack nuance, framing claims as self-evident (or universal), and thereby reflecting a strong claim to absoluteness. For dimension 6 (*emotionality*), bigrams reflect emotional states including urgency, concern, and fear. Calls to action include *teilen teilen*, expressions of subjectivity appear in *ich hoffe* or *ich glaube*. Direct address or inclusion of the audience is present in *du bist, lasst euch*, and *wir alle*.

8 Discussion & Conclusion

Drivel can be understood as encompassing two distinct aspects: (i) the narrative content, which involves the presentation and framing of specific topics; and (ii) topic-independent linguistic features that contribute to the overall drivel-like quality of texts irrespective of its narrative content. While much prior research was focused on the former, the present study has concentrated on the latter. By analysing a data set containing a large proportion of conspiracist and "drivelly" posts, we have intensionally defined and extensionally annotated six dimensions capturing various textual characteristics, notably its *distance from reality, linguistic and argumentative peculiarities, claim to absoluteness, suggestiveness, oversimplification*, and *emotionality*. In summary, our analyses demonstrate that content-agnostic features of drivel can reliably be annotated, and, to a considerable extent, predicted from textual patterns. We report substantial IAA for all dimensions except *linguistic and argumentative peculiarities* and *suggestiveness*.

Linear regression analyses revealed that the overall drivel-like quality of posts is primarily driven by *distance from reality* and *claim to absoluteness*, with *emotionality* contributing less in both the full and a reduced model.

Recognising that drivel exists on a continuum, we employed ordinal regression to model these dimensions, demonstrating that several can be reliably predicted from textual features without reference to external world knowledge. In particular, *linguistic and argumentative peculiarities, claim to absoluteness*, and *emotionality* can be predicted effectively from *n*-grams. Feature weight analyses of bigrams revealed that high-scoring posts frequently employ both content-specific and generic linguistic patterns, reflecting rhetorical, stylistic, and emotionally evocative choices. Overall, these results provide a characterisation of the linguistic and structural markers of drivel, offering insights into generic conspiratorial patterns of language complementing work on narrative content.

Future research will aim to refine the dimensions based on the insights obtained here, to extend the manually annotated data set, and to improve automatic classifiers and recognisers for each dimension. We will also extend the feature extraction methodology presented here, based on linguistic insights on the one hand, such as sequences of parts of speech, and automatic methods, such as vector embeddings, on the other. Overall, we hope to contribute to a better understanding of drivel in general, and enhance the capacity to detect and analyse it in large-scale social media corpora.

Funding This research has been partially funded by the German Research Foundation (DFG), project no. 466328567.

Funding Open Access funding enabled and organized by Projekt DEAL.

References

1. Douglas KM, Uscinski JE, Sutton RM, Cichocka A, Nefes T, Ang CS, Deravi F (2019) Understanding conspiracy theories. Polit Psychol 40(S1):7. https://doi.org/10.1111/pops.12568
2. Heinrich P, Blombach A, Doan Dang BM, Zilio L, Havenstein L, Dykes N, Evert S, Schäfer F (2024) Automatic identification of COVID-19-related conspiracy narratives in German Telegram channels and chats. In: Calzolari N, Kan MY, Hoste V, Lenci A, Sakti S, Xue N (eds) Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024). ELRA and ICCL, Torino, Italia, pp 1932–1943
3. Blombach, A., Doan Dang, B.M., Evert, S., Fuchs, T., Heinrich, P., Kalashnikova, O., Unjum, N.: Narrlangen at SemEval-2025 task 10: Comparing (mostly) simple multilingual approaches to narrative classification. In: Proceedings of the 19th International Workshop on Semantic Evaluation. SemEval 2025, Vienna, Austria (2025)
4. Piskorski, J., Mahmoud, T., Nikolaidis, N., Campos, R., Jorge, A., Dimitrov, D., Silvano, P., Yangarber, R., Sharma, S., Chakraborty, T., Guimarães, N., Sartori, E., Stefanovitch, N., Xie, Z., Nakov, P., Da San Martino, G.: SemEval-2025 task 10: Multilingual characterization and extraction of narratives from online news. In: Proceedings of the 19th International Workshop on Semantic Evaluation. SemEval 2025, Vienna, Austria (2025)
5. Singh I, Scarton C, Bontcheva K (2025) GateNLP at SemEval-2025 task 10: Hierarchical three-step prompting for multilingual narrative classification. In: Rosenthal S, Rosá A, Ghosh D, Zampieri M (eds) Proceedings of the 19th International Workshop on Semantic Evaluation (SemEval-2025). Association for Computational Linguistics, Vienna, Austria, pp 148–154. https://aclanthology.org/2025.semeval-1.21/
6. Bourdieu P (1991) Language and Symbolic Power. Harvard University Press, Cambridge, MA
7. Frankfurt HG (2005) On Bullshit. Princeton University Press, Princeton, NJ
8. Blombach A, Evert S, Havenstein L, Heinrich P (2025) Dimensions of drivel in German telegram posts: Manual annotation and predictive power. In: Proceedings of CMC 2025, pp 50–54
9. Lamberty P, Holnburger J, Goedeke Tort M (2022) Das Protestpotential während der COVID-19-Pandemie. https://cemas.io/blog/protestpotential/
10. Holnburger J, Goedeke Tort M, Lamberty P (2022) Q vadis? Zur Verbreitung von QAnon im deutschsprachigen Raum. https://cemas.io/publikationen/q-vadis-zur-verbreitung-von-qanon-im-deutschsprachigen-raum/
11. Warnke L, Maier AL, Gilbert DU (2024) Social media platforms' responses to COVID-19-related mis- and disinformation: the insufficiency of self-governance. J Manag Gov 28:1079–1115. https://doi.org/10.1007/s10997-023-09694-5
12. Schäfer F, Evert S, Heinrich P (2017) Japan's 2014 General Election: Political Bots, Right-Wing Internet Activism and PM Abe Shinzō's Hidden Nationalist Agenda. Big Data 5(4):294–309
13. Institut für Demoskopie Allensbach: Politischer Radikalismus und die Neigung zu Verschwörungstheorien (2022)
14. Kuhn SAK, Lieb R, Freeman D, Andreou C, Zander-Schellenberg T (2021) Coronavirus conspiracy beliefs in the German-speaking general population: endorsement rates and links to reasoning biases and paranoia. J Mil Gov Couns. https://doi.org/10.1017/S0033291721001124
15. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33(1):159–174
16. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: Machine learning in Python. Journal of Machine Learning Research 12:2825–2830
17. Pedregosa-Izquierdo F (2015) Feature extraction and supervised learning on fMRI : from practice to theory. Theses, Université Pierre et Marie Curie - Paris VI. https://theses.hal.science/tel-01100921
18. Rennie JDM, Srebro N (2005) Loss functions for preference levels: Regression with discrete ordered labels. In: Proceedings of the IJCAI Multidisciplinary Workshop on Advances in Preference Handling
19. Baker P (2006) Using Corpora in Discourse Analysis. Continuum, London