“You are what you Tweet!” - Analysing Pro-Anorexia and Pro-Recovery Messages On Twitter Using Transformer-Based Text Mining Applications

MSc. Thesis

Julius Heinke

University of Twente

Faculty of Behavioural Management Sciences (BMS)

Department of Technology, Human and Institutional Behaviour (HIB)

Section of Psychology and Health Technology

Supervision and Examination Committee

Dr. P.M. ten Klooster

Dr. M. Nieuwenhuis

August 27, 2023

UNIVERSITY OF TWENTE.

Abstract

Background & Objective. "Pro-anorexia" (PA) communities that encourage disordered eating are a common phenomenon on social media sites like *Twitter*. In response, "pro-recovery" (PR) communities that postulate a recovery-oriented perspective emerged. Past research has struggled with the vast quantity of social media content, limiting their insights into these communities. To better understand PA/PR community content and dynamics, this study utilized modern transformer-based text mining methods and metadata to analyse all English PA/PR-related tweets posted until February 1, 2023.

Method. A total of 288.773 PA and 17.839 PR tweets were scraped using a specified set of hashtags, and analysed on their metadata, sentiment, and discussed topics using novel *roBERTa-based-sentiment* and *BERTopic* analyses in *Python*. Social metrics (likes, retweets etc.), overall findings, and changes over time in popularity, topics, and sentiment were analysed and compared.

Results. More than 16 times as many PA than PR tweets were found. PA tweeters had more likes, retweets, quotes and replies. PR users posted more tweets, followed more accounts, and engaged in broader content. PA sentiment was both more often positive and negative. PR sentiment was more neutral and more stable. For both communities, higher tweet sentiment coincided with greater tweet frequency in one community and lower sentiment and fewer tweets in the other. PA topics were focussed on anorexia and over time shifted from inspiration/connection to no current primary topic. PR had a broader focus and over time showed a shift from discussing eating habits to media content.

Conclusion. Based on metadata and topics, the PA community appeared to show a stronger focus on social interaction within the community and anorexia-related topics, while also being more secluded from other communities. Sentiment and topics indicate that PA tweets tended to more often use polarising language and weight loss encouraging topics. The PR community appears more neutral, with a larger focus outside AN. Future research is needed to verify these findings and test methods of intervention.

Contents

Introduction

Every year, more than 3.3 million "healthy live years" worldwide are lost to eating disorders (EDs) (van Hoeken & Hoek, 2020). In fact, more than 9% of the world population are impacted by EDs (Arcelus et al., 2011). While recent years have brought an increasing recognition and awareness of the morbidity and mortality caused by EDs, a great need for efficacious interventions remains (Arcelus et al., 2011). Of all the EDs, Anorexia Nervosa (AN) is not only the most deadly, but also the hardest to treat, creating the greatest need for effective interventions (Arcelus et al., 2011; Schmidt et al., 2018). AN is defined as "an eating disorder characterized by a significantly low body weight [...] generally obtained by means of severely restricted food intake" (DeWitt & Attia, 2017, p. 13). Individuals with anorexia "display an intense fear of weight gain and a distorted self-evaluation of his or her weight and shape" (DeWitt & Attia, 2017, p. 13). These symptoms cause great physical and mental suffering in affected individuals. Some studies suggest more than 20% of AN sufferers to die from the condition (Sullivan, 1995). Due to the COVID-19 pandemic, many AN sufferers were disconnected from support systems, causing admission and readmission rates to increase more than eight times (Matthews et al., 2021).

Anorexia & Media

Anorexia nervosa is a highly stigmatized disorder (Rich, 2006). AN sufferers report feeling like their condition is misunderstood by peers, family, friends, and even healthcare professionals (Rich, 2006). As a result, individuals suffering from AN often try to conceal their symptoms. In fact, many never even discuss their experiences with professionals, making it hard to initiate treatment for them (Swanson et al., 2011). Searching for understanding and comfort, many people with AN build relationships with other sufferers and engage themselves in media representing a "thin ideal" (Rich, 2006; Sukunesan et al., 2021). It has long been established that traditional media and its representation of a "thin ideal" constitute a significant risk and maintenance factor of anorexia by increasing body dissatisfaction (Ferguson et al., 2014; Grabe et al., 2008).

However, the rise of social media (SM) in recent years has led to substantial changes to the media landscape, largely replacing traditional media (Sukunesan et al., 2021). About 84% of young adults in western countries such as the US, which are most at risk for AN, use SM (Hosokawa et al., 2023; Silén & Keski-Rahkonen, 2022). Most recently, the COVID-19 pandemic led to an additional 20% increase of total time spent on SM (Hall & Liu, 2022). Jordan et al. (2021) also found significant increases in eating-disorder-related SM content.

The unique ability to converse and interact completely anonymously on SM strongly coincides with the secretive nature of AN, and makes seeking social support through SM even more attractive (Arseniev-Koehler et al., 2016). Further, SM allows real-time peer interactions (Sidani et al., 2016). This ability to exchange experiences or worldviews, and more directly bond with others in very similar life situations creates strong social fields, organized into "communities" (Giordani & Silva, 2021). However, the emergence of SM also shifted these previously hidden communities to the public, where they could now be investigated. This potentially allows new insights and understanding of AN communities that could then inform future interventions. There are a broad range of online communities, and sub-communities, that concern themselves with AN, organized by shared values, discussed topics and social relations (Wang et al., 2017). The two largest opposing communities are the so-called *pro-anorexia* and *pro-recovery* communities.

Pro-Anorexia vs Pro-Recovery Communities

"Pro-anorexia" or "Pro-ana" (PA) communities revolve around encouraging and promoting anorexic behaviour (Branley & Covey, 2017). Rather than acknowledging AN as a serious psychological disorder that requires treatment, it is presented as a lifestyle choice or even a virtue (Branley & Covey, 2017). A choice that is required to be part of this community (Giordani & Silva, 2021). Members of PA communities continuously promote thinness and glorify the ideal of a thin body (Sukunesan et al., 2021). In their interactions, community members share their weight loss progress and struggles and receive compliments and/or advice (Mento et al., 2021; Sukunesan et al., 2021). They also frequently exchange information on how to best lose weight, such as restrictive dieting, extreme exercise, or so called "thinspiration" (sharing motivational images of thin bodies) (Borzekowski et al., 2010; Chancellor, Pater, et al., 2016). As a result, losing weight is not only normalized, but has become the central tenet of this community (Branley & Covey, 2017).

With no existing offline equivalents, the PA community is unique and receives a great deal of societal, governmental, and organizational attention for their negative effects (Sukunesan et al., 2021; Yom-Tov et al., 2012). Past research established correlations between the engagement in PA content and worsening symptoms of anorexia due to the prevention of help-seeking, reinforced disordered eating, and the portraying of a wrong sense of "support"(Harper et al., 2008; Rouleau & von Ranson, 2011; Wilson et al., 2006). Giordani and Silva (2021) even warn about the clinical consequences of PA communities functioning as "anti-treatment and anti-recovery devices" (Giordani & Silva, 2021, p. 5294).

In contrast to PA communities, pro-recovery (PR) communities focus on increasing awareness and promoting recovery from AN. Here, information and support regarding treatment seeking and AN recovery can be found (Branley & Covey, 2017). The offered guidance and support of PR communities can help individuals in recovering from AN (Branley & Covey, 2017). By increasing the users' knowledge, behaviour intent, and their awareness, SM poses a powerful platform for promoting health behaviours in AN sufferers (Chung et al., 2021). These online support communities are especially valuable when considering that sufferers rarely seek professional help (Cachelin & Striegel-Moore, 2006).

The Importance of Understanding AN Communities

Pro-ana censorship. In response to the potential dangers of PA communities, it has previously been proposed to censor this content on the internet. Suggestions for censorship range from intervening in the search results, or adding warning labels when anorexia-related content is accessed, to banning PA content altogether (Yom-Tov et al., 2012). The governments of France and the UK have even enacted legislation for restricting PA content (Assemblée nationale, n.d.; DCMS, n.d.). Wary of these developments, major

SM sites such as Facebook, Instagram, Tumblr, and Pinterest have implemented policies for blocking PA related hashtags and content (Arseniev-Koehler et al., 2016; BBC News, 2012; Instagram, n.d.; Pinterest, n.d.).

Negative impact of censorship. However, the decision to censor PA content is highly controversial. Despite increasing regulatory pressure and social stigma, PA communities have not shrunk in size (Casilli et al., 2013). Network analyses by scholars such as Casilli et al. (2013) revealed that PA communities instead turn inwards, and progressively seclude themselves from other communities. PA community members continue exchanging messages, links, and images in gate-kept communities but now exclude any outside information sources, including health information, and awareness campaigns (Casilli et al., 2013). Consequently, health information or awareness campaigns, offered to help people suffering from AN, are made much less likely to reach their targets (Casilli et al., 2013). Further, AN sufferers may lose existing beneficial social connections, negatively impacting their well-being (Branley & Covey, 2017).

Other parts of the PA community have also been observed to obscure hashtags such as `#thinspo` to `#thynspo`, and/or display disclaimers that a community space is not PA, when it actually is (Cobb, 2017). This way, they effectively circumvent any of the censorship measures attempted so far (Cobb, 2017). Such denial and disguise bears the danger for PA to be further normalized, blurring the lines between PA and "healthy thinness" (Cobb, 2017). An increasingly obscured and disguised community is much harder to reach through intervention, and might not be recognized as unhealthy. The fact that censorship may not simply be ineffective, but can be considered actively harmful for people suffering from AN stresses the importance of truly understanding online AN communities, before introducing any interventions (Branley & Covey, 2017).

Previous Research on Anorexia Communities

As aforementioned, PA community members aim to support each other in their condition. Park et al. (2022) has found such support to predominantly take the shape of emotional support rather than factual/informational support. In comparison to other ED communities, PA online communities are also much more rigid in their beliefs while denying any negative aspects of AN (Lai et al., 2021). Further, they show a greater focus on actively sharing content and exhibit greater social adherence, with more "likes" per "view" than any other ED community (Lai et al., 2021). Being such a tight-knit, active community that is focused on illness progression, rather than treatment, further exacerbates the potential for harm of PA communities.

Despite this, PA communities have also been shown to offer room for PR communities to emerge (Park et al., 2022). Considering some of the previously detailed benefits of PR communities for AN sufferers, this is a promising finding. Nevertheless, little is known about the exact impact of PR communities on individuals with AN, and it is unclear how the PR community interacts with the PA community. Further, it needs to be understood in which ways these communities differ, as this way PR communities could be made more attractive for individuals that would otherwise turn to PA communities. Previous research has mostly investigated PA content while disregarding PR communities (Branley & Covey, 2017). This provides an incomplete picture, since anorexia communities are very dynamic, with users changing from PA to PR communities and vice versa (Yom-Tov et al., 2012). Further, the fast-paced nature of online media creates a need to (re-)interpret any existent findings from the perspective of current SM practices.

To address the limitations of previous research, it would be insightful to investigate and compare aspects such as discussed topics, sentiments, and interactions within both PA and PR communities. This way, a more complete understanding of AN communities, as they currently manifest, overlap, and differ can be generated.

Using Twitter to better understand AN. *Twitter*, a major social networking service that became popular for its microblogging capability, is the largest SM platform that has not introduced censoring of PA communities (Arseniev-Koehler et al., 2016; Branley & Covey, 2017; Sukunesan et al., 2021). This makes it a valuable space for investigating both PA and PR community content and dynamics as they naturally occur. Twitter users can create their own profiles with usernames called *handles* and express themselves through *tweets*, publicly posted messages of up to 280 characters in length (Sukunesan et al., 2021). Given the anonymity of Twitter, users can also be expected to disclose information much more freely, offering deeper, less biased insights into AN communities. With a large user-base that are, in large parts, in the age and geographic background most at-risk for AN, it also offers an important target for further stimulating AN recovery (Statista, n.d.).

The Potential of Text Mining

As all tweets are posted publicly, well over a decade's worth of online interactions of AN communities are available to scientific research. Nevertheless, it remains difficult for researchers to manually process this into a useful dataset and perform analyses. As a result, previous studies on PA and PR communities were constrained by the amount of data they considered in their manually performed analyses (Arseniev-Koehler et al., 2016). Further, the sheer amount of data has made it hard to perform meaningful content analyses, limiting the insights about increasingly complex interactions of AN communities (Branley & Covey, 2017).

A promising tool for overcoming the aforementioned limitations of previous research is *text mining*. Text mining refers to the "extraction of information and patterns that are implicit, previously unknown, and potentially valuable [...] from immense unstructured textual data" (Hassani et al., 2020, p. 2). This is achieved by using computer algorithms that are, usually, trained through machine learning to operate fully automatically (Hassani et al., 2020). Two frequently employed text mining applications are topic modelling and sentiment analysis. With topic modelling, it is possible to automatically extract the topic(s) of a text (Albalawi et al., 2020). Sentiment analyses help in automatically determining potential underlying attitudes, views, feelings, and opinions of a text by attributing a "sentiment score" (Yadav & Vishwakarma, 2020).

Text mining could offer an effective method for extracting meaning such as discussed topics, or sentiments from an unlimited amount of tweets. Such a big-data approach offers a variety of advantages. For one, larger datasets introduce a level of objectivity by abstaining from researchers hand-picking a select few tweets. Further, by considering much more of the available data, a broader picture of SM practices can be painted. Using historic data, this also allows for analyses on developments over time. Recent developments of so-called transformer-based text-mining applications further facilitated the capabilities of text-mining to more accurately consider figures of speech like irony and sarcasm, abbreviations and emojis.

Prior studies have already made effective use of text mining methodology to investigate parts of AN communities (Tiggemann et al., 2018; Zhou et al., 2020). Nevertheless, only Fettach and Benhiba (2019) have made use of topic modelling and sentiment analyses, by investigating both communities on *Reddit*, in which they found an organisation into multiple sub-communities, and polarisation of sentiment in PA communities (Fettach & Benhiba, 2019). This constitutes a research gap of studies utilising the aforementioned advantages of investigating PA and PR communities on Twitter while utilising novel transformer-based text mining applications.

Research Objective & Research Questions

Attempting to fill the aforementioned research gap, this study utilizes novel transformer-based topic modelling and sentiment analyses to examine the discussed topics and sentiments of PA and PR communities on Twitter, as well as how they developed over time. Furthermore, metadata such as likes, retweets, and followers are used to explore the interactional dynamics, popularity of content and users for further facilitating the understanding of both communities.

The subsequent research questions (RQs) are as follows:

RQ1: *How do the pro-anorexia and pro-recovery communities differ in terms of social metrics such as tweet frequency and number of likes/retweets/followers?*

RQ2: *How does the sentiment of pro-anorexia and pro-recovery tweets differ?*

RQ3: *How did the sentiments of pro-anorexia and pro-recovery tweets develop over time?*

RQ4: *How do the discussed topics of pro-anorexia and pro-recovery tweets differ?*

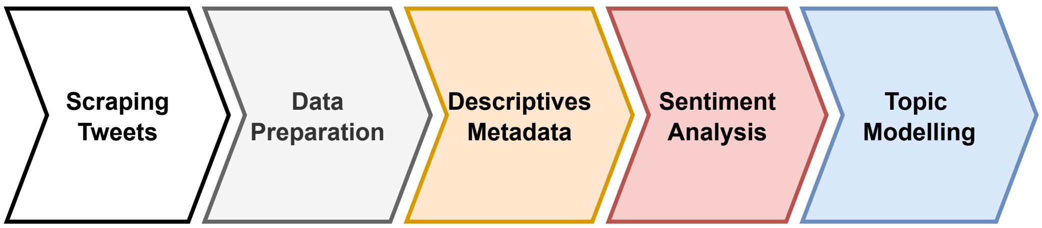
RQ5: *How did the discussed topics of pro-anorexia and pro-recovery tweets develop over time?*

Methods

To answer the aforementioned research questions, several steps were taken to collect the relevant data and conduct the subsequent analyses. The sequence of the main steps is visualized in Figure 1. Beginning with scraping the tweets, they were then preprocessed and analysed on the available metadata, their sentiment, and topics.

Figure 1

Overview of the steps undertaken in this research.



Data Gathering

Tweet Inclusion Criteria

The first step of data gathering involved determining which tweets to include. To tag their posts, Twitter users use so-called “hashtags”, a crowdsource tagging system, by adding `#` followed by the tag-word to their tweet (Akundi et al., 2018). Twitter users have an interest in correctly tagging their tweets, as they want to reach their respective community. Further, as Akundi et al. (2018) point out, hashtags are also used to engage in ongoing conversations. Consequently, hashtags are more precise than simple keywords in determining a tweet’s content, and many previous research studies have used them as an inclusion criterion (Akundi et al., 2018; Branley & Covey, 2017; Sukunesan et al., 2021).

Since most communities use multiple hashtags, first the main hashtags of PA and PR communities were determined. This was done by first scraping all tweets that contained the main hashtags `#proana` and `#prorecovery`. Then, it was determined which other AN-related hashtags were most frequently used in conjunction with these. The most frequently used PA and PR community specific hashtags, (e.g. not like `#sport`), were then selected as inclusion criteria for final scraping. More general hashtags such as `#anatwt`, which are not clearly associated with one of the two communities, or hashtags that are used by outside communities (e.g.: `#beatana` - also used in sports), were excluded. Alternative writing styles of the original hashtags were also included (e.g.: `pr0ana`). Table 1 shows the final list of hashtags for each community.

Table 1

Most relevant Twitter hashtags by community.

Pro-anorexia	**Pro-recovery**
#proana; #pr0ana	#prorecovery
#anasister; #anasisters	#anorexiarecovery; #anarecovery
#anabuddy; #4nabuddy	#anorexiafighter; #anafighter
#meanspo; #meansp0	#anorexiawarrior; #anawarrior
#sweetspo; #sweetsp0	#anorexiasoldier; #anasoldier
#bonespo; #bonesp0	#beatanorexia
#thinspiration; #thinspotwt; #thinspo;	#eatingdisorderrecovery; #edrecovery
#thinsp0; #th1nspo; #th1nsp0; #thynspo	#edfighter; #edwarrior; #edsoldier

Scraping Tweets

As a second step, all relevant tweets were collected into a dataset through a process called *web scraping*. Essentially, scraping describes the process of systematically extracting and combining content from the internet (Glez-Peña et al., 2014). To scrape the tweets, the Python library *tweepy* by Harmon et al. (n.d.) was used to access Twitter's academic research Application Programming Interface (API). The Python script used in this, and all other analyses, is accessible on GitHub (Appendix A). On February third 2023, all English tweets containing one of the hashtags in Table 1 posted from the founding of Twitter until February first 2023 were scraped. In addition to the tweet itself, additional metadata, such as usernames, profile descriptions, post time, and number of likes, retweets, followers, and following were collected. All the scraped data was stored into a comma-separated values

(CSV) file using the Python library NumPy (Harris et al., 2020). All datasets are available on GitHub (Appendix A). In total, 383.864 tweets were scraped. Of these, 288.773 tweets used at least one PA hashtag, and 95.091 tweets used at least one PR hashtag.

Data Preprocessing

Of the PR hashtags, some are not exclusively used in AN contexts. Namely, `#prorecovery`, `#eatingdisorderrecovery`, `#edrecovery`, `#edfighter`, `#edwarrior`, and `#edsoldier`, were also used by other ED communities. Nevertheless, as they contain a number of anorexia-related PR tweets, these need to be synthesized from the scraped tweets. Therefore, all tweets that did not contain any of the other PR hashtags or the keywords *anorexia* or *ana* were excluded from the dataset. In total, 77.252 tweets were excluded. The final amount of tweets of the PR community was 17.839. As the PA dataset only contained tweets with hashtags clearly related to AN, no tweets were excluded.

To prepare the tweets for subsequent sentiment analyses, they had to be preprocessed. More specifically, *roBERTa-based-sentiment*, the model used for extracting the sentiment (elaborated on in the subsequent sections), was trained on a dataset of tweets where user mentions, as well as hyperlinks, were processed to be uniform and anonymous (Loureiro et al., 2022). Thus, all user mentions were replaced with "*@user*", and all hyperlinks with "*http*". This assured that the semantic structure remained the same, while preventing user mentions and links to influence the sentiment of the tweet itself. Similar preprocessing was tested for topic modelling, but it was found that this substantially reduced the quality of the extracted topics for BERTopic-based topic modelling. Thus, no pre-processing was performed for the topic modelling.

Descriptive Metadata Investigations

To investigate how the PA and PR communities differ in terms of social metrics such as tweet frequency and the number of likes/retweets/followers (RQ1), several descriptive analyses were conducted using the metadata scraped for every tweet. Using the date- and timestamp each tweet was posted, the frequency and volume of tweets over time was explored and compared between the two communities using *Matplotlib* and *NumPy* (Harris et al., 2020; Hunter, 2007). After this, it was investigated how likes, retweets, replies, follower and following counts differ between the communities. As the mean shows the average engagement, a higher mean count on these metrics was considered to indicate greater engagement and interaction within the community. The max count indicates the largest value and, together with the standard deviation, higher values hint towards larger, influential actors within the community. Retweets and replies were counted towards the total number of tweets of each user. The *lists* describe the type of content a user engages in, the more lists a user is on, the more diverse the content they engage in is. These analyses were conducted using the Python library *SciPy* (Virtanen et al., 2020). To assess the overlap and interaction between the PA and PR community, it was compared how many tweets occur in both datasets and how many users posted in both communities.

Text Mining Analyses

In the third step, the main analyses were conducted using the previously prepared dataset of tweets. Each tweet contains a written text of up to 280 characters (140 until 2017). To make sense of these texts, it is essential to consider several factors, such as the meaning of words, vocabulary, syntax, and grammar. In the past, this meant that only humans were capable of processing language. This changed with the development of *Natural Language Processing (NLP)* which allowed the emergence of software that can make sense of human language (Chowdhary, 2020). Systems based on NLP software are capable of quantitatively analysing as well as organizing text information autonomously . NLP does not only allow for the interpretation of large amounts of texts, it may also introduce a greater level of robustness than human researchers by continuously relying on the same principles (Garman et al., 2018). In this research, the two NLP-based applications of topic modelling and sentiment analysis were applied to analyse the text of all tweets.

Sentiment Analysis

First, a *Sentiment Analysis* of the scraped tweets was conducted. In this context, *sentiment* refers to the underlying attitudes, views, feelings, and opinions expressed in the text of a tweet (Yadav & Vishwakarma, 2020). A sentiment analysis describes the process of automatically determining these underlying features (Shelar & Huang, 2018; Yadav & Vishwakarma, 2020). In the context of this study, sentiment analysis provides important insights into the polarity of tweets posted in the PA and PR community and enables inferring differences between the two (Shelar & Huang, 2018).

The most recent version of the *Twitter-roBERTa-base for Sentiment Analysis*, developed by Loureiro et al. (2022), was used to identify the sentiment of each tweet. *roBERTa* stands for "*robustly optimized BERT approach*" and is a further optimized version of *Bidirectional Encoder Representations from Transformers* (*BERT*), an open source machine learning framework for NLP developed by Devlin et al. (2019). RoBERTa was developed to improve task performance for pre-trained models (Liu et al., 2019). The Twitter-roBERTa-base sentiment model was trained using 124 million tweets from January 2018 to December 2021, and fine-tuned for sentiment analysis with the *TweetEval* benchmark (Loureiro et al., 2022).

Traditional sentiment analyses rely on a lexical approach, where every word is given a certain score, based on which the sentiment of a tweet is calculated (Elbagir & Yang, 2019). While this is a simple and effective method, its accuracy is strongly reduced when applied to the more complex texts found on SM, such as tweets. Tweets of anorexia communities often contain irony, sarcasm, emojis, and slang that render lexical approaches ineffective. The roBERTa-base model considers the whole semantic structure and content of a tweet for determining its sentiment. This made it possible to detect and consider figures of speech like irony or sarcasm. Further, as this model was trained exclusively on tweets, it is familiar with their short length, emojis, abbreviations, and common slang.

Application of Twitter-roBERTa-base sentiment in Python. The model used for sentiment analyses was accessed in Python through the *TweetNLP* API developed by Camacho-Collados et al. (2022). Using the preprocessed text, TweetNLP was utilized to calculate the sentiment of each tweet. These sentiments were provided as a percentage of how likely it is that the tweet is negative, neutral, or positive. Additionally, the most likely sentiment is indicated categorically. To enable subsequent analyses, the sentiment of each tweet was then transformed to a continuous scale from -1 (negative) to 1 (positive), with 0 being neutral. The overall sentiment was then compared by calculating the mean and standard deviation of PA and PR tweet sentiment. These scores were used to compare how the sentiment of PA and PR tweets differ (RQ2). Further, several graphs of the sentiment of both communities over time were created to explore how the sentiments of PA and PR tweets developed over time (RQ3).

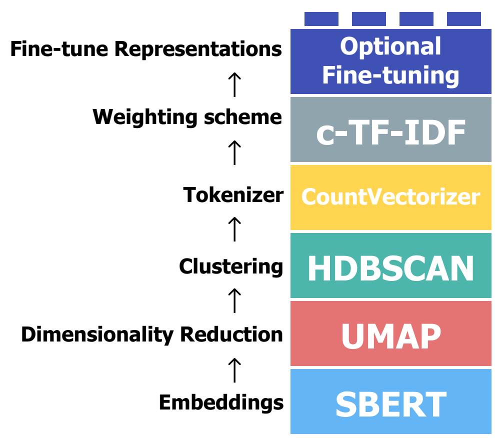
Topic Modelling

Latent themes of both the PA and PR community were identified through *topic modelling*. Topic modelling is a term used for a "group of algorithms that reveal, discover, and annotate thematic structure" of texts (Kherwa & Bansal, 2018, p. 2). Topic modelling algorithms allow the determination of hidden semantics in texts and cluster the themes that emerge into topics (Kherwa & Bansal, 2018). These automatically generated topics can then be used to derive meaning from a text. Contemporary literature indicates that previous, go-to, algorithms perform considerably worse on tweets due to their unique characteristics (unstructured, less than 280 character, use of sarcasm/irony) (Egger & Yu, 2022; Jonsson & Stolee, n.d.). Further, conventional models describe documents as a bag-of-words, completely disregarding the semantic relationships between words in tweets, limiting the accuracy of topics (Grootendorst, 2022).

BERTopic. A promising alternative is a newly emerging topic modelling technique called *BERTopic* (Grootendorst, 2022). In a review of all major established and upcoming techniques by Egger and Yu (2022), BERTopic has shown to perform well on all aspects of topic modelling, whereas most other techniques typically excel at one aspect only. Especially for tweets, BERTopic outperforms any of the aforementioned go-to techniques. Just like roBERTa based sentiment, BERTopic was built using BERT. BERTopic is built to be flexibly adapted and offers interactive visualization, allowing for deeper insights into each topic (Egger & Yu, 2022). More specifically, BERTopic is built on three main pillars, the ability to: (1) build your topic model, (2) explore various techniques (in training and post-hoc), (3) create topic visualizations (Grootendorst, 2022).

Figure 2

Illustration of the BERTopic Algorithm Pipeline used in this research.



Build your topic model. In this study, the default configuration (Figure 2) was adapted to perform best on the data used (e.g. English, short texts, etc.). The embedding extraction, using *SBERT*, marked the first step of the pipeline used in this research. A considerable advantage of SBERT is that it considers the semantic relationships of the words in a tweet, capturing a more complete representation of topics as a result (Grootendorst, 2022). To create accurate embeddings with transformers, manual pre-processing of data is not only not required, but actually advised against. By considering the whole tweet, without removing anything prior to the analyses, transformers consider their full context to create the embeddings. This does not only simplify the workflow, but also reduces the chance of human error or bias in manual preprocessing (Egger & Yu, 2022). *UMAP* was used to reduce the dimensionality of the extracted embeddings. UMAP is a stochastic algorithm, meaning that the results from BERTopic may differ when the same model is run multiple times (Grootendorst, 2022). This was utilized to improve result accuracy by repeatedly running the dimensionality reduction until the most interpretable topics were found. *HDBSCAN* was used for clustering the embeddings. An advantage of HDBSCAN is that it automatically identifies outliers, so any tweet that does not contain a relevant topic is excluded as an outlier. In the next step, *CountVectorizer* was used to generate the topics from the previously reduced and clustered embeddings. CountVectorizer uses a bag-of-words approach on a cluster level to extract higher-level topics, without making any assumptions on the clusters themselves. In a last step of the pipeline, *class-based term frequency-inverse document frequency* (*c-TF-IDF*), as well as custom representations trained through *ChatGPT 3.5* were used to generate coherent topic labels and representations.

Explore Various Techniques. In addition to the modularity of the algorithm, BERTopic is also flexible in the techniques that can be used both for training the model, and representing the findings. In this research, unsupervised topic modelling was used to extract the topics from the tweets. From the available post-hoc techniques, hierarchical, and dynamic topic modelling (DTM) were utilized (Grootendorst, 2022). Hierarchical topic modelling allows an overview of how topics relate to each other. DTM overcomes the static nature of traditional topic modelling techniques, by enabling topics to be modelled over time for specific timestamps. Considering the fast-paced nature, and long timeframe investigated in this study, DTM was deemed especially useful. It was expected that prevalent topics among both communities changed significantly throughout the 16 years.

Create Topic Visualizations. With no right or wrong results, it is challenging to properly interpret the results of topic modelling. Visualizations offer an important means to both understand the topics and their relations, and to present them. BERTopic offers many useful ways of visualizing extracted topics (Grootendorst, 2022). For example, it is possible to create a two-dimensional representation of all topics in an *intertopic distance map*. This gives an immediate overview of similarities in the semantic structure and occurrence of all topics. All BERTopic visualizations are interactive, enabling a deeper level of engagement and greater detail. It is also possible to create interactive hierarchies of topics to further understand the relationships of the topics. Moreover, several visualizations allow for deeper investigation of the relevance of each word for its respective topic. For DTM, BERTopic also allows the representation of topic frequency over time. These interactive visualizations were used for investigating and presenting the topics discussed in both communities.

Application of BERTopic in Python. To employ BERTopic in Python, the library *bertopic*, by Grootendorst (2022) was used. To achieve the most interpretable results, some parameters of the utilized pipeline were tuned. This involved an iterative process of running different settings until the best possible topic representations were extracted. First, the embeddings were extracted from the tweets using the SBERT model *all-mpnet-base-v2* rather than the default *all-MiniLM-L6-v2*. This sentence-transformer is about five times more computationally intensive, but allows for the most accurate embedding extractions with BERTopic (Reimers & Gurevych, 2019). Using the extracted embeddings, the topics were identified using the pipeline detailed on the left of Figure 2. With the large number of documents, the initial number of topics exceeded 2000. The following steps were undertaken to reduce the number of topics while retaining as much information and quality of topic representation as possible.

First, the BERTopic parameter `min_topic_size` was increased from its default value of *10* to *15* for PR and *180* for PA topic modelling. This parameter determines the minimum number of tweets that needed to contain this topic. Increasing it leads to fewer, more prominent topics being created. The parameter values lie in the sweet spot of reducing the number of generated topics while not decreasing the quality of these topics. Further, the UMAP parameter `n_neighbours` was increased from its default value of *15* to *40* for extracting the PR and PA topics. Increasing this parameter changes how UMAP balances the data structure to consider a more global overview of tweet topics (McInnes et al., 2018). This not only reduced the number of final topics, but also improved the quality of the extracted topics by creating larger clusters (Grootendorst, 2022).

While the aforementioned steps already decreased the number of topics generated, the most essential step involved the actual topic reduction by merging similar topics. In most other topic modelling techniques, topics are merged until a specified number of topics is reached. This approach is flawed in that topics are merged regardless of whether they are actually similar, but simply because they are *most* similar. In contract, BERTopic allows for *automatic* topic reduction by use of *HDBSCAN*. Such automatic topic reduction only merges those topics that actually belong together, based on their content, as determined by their semantic and thematic relatedness. In this step, the number of topics were reduced from roughly 180 topics to 85 in the PR, and 93 in the PA dataset. As the text was not pre-processed in any way, stop words frequently made it into the topic representations. To remove these stop words, without limiting the effectiveness of the utilized transformer models, *CountVectorizer* was used to remove them only after having generated the embeddings and clustered the documents. The rest of the parameters were left at their default values.

Using these parameters, the pipeline was run repeatedly until a satisfactory topic representation was achieved. After this, initial custom labels for the topics were created using ChatGPT. More specifically, the ChatGPT API was accessed and prompted, through an integrated wrapper within BERTopic, to automatically generate topic labels one-by-one for each of the 178 total topics. The prompt used in this research can be found in Appendix D. Topic labels were created based on the five most relevant topic terms in the initial representation, as well as the three most representative tweets of each topic as determined by their loading onto the topic. This way, labels could be generated that not only took the most frequently appearing topic terms into account, but also the context in which they appear, by looking at the tweets that most closely resemble these topics. As a last step, final topic labels were created using the labels created by ChatGPT, the most frequent topic terms, the most representable documents and by use of visualizations of topic term relevancy and overall clustering of the topics to manually investigate more ambiguous topics. The visualization of topics over time using dynamic topic modelling was created with *nr_bins*, the number of bins into which the timeframe was split, set to *60*.

Results

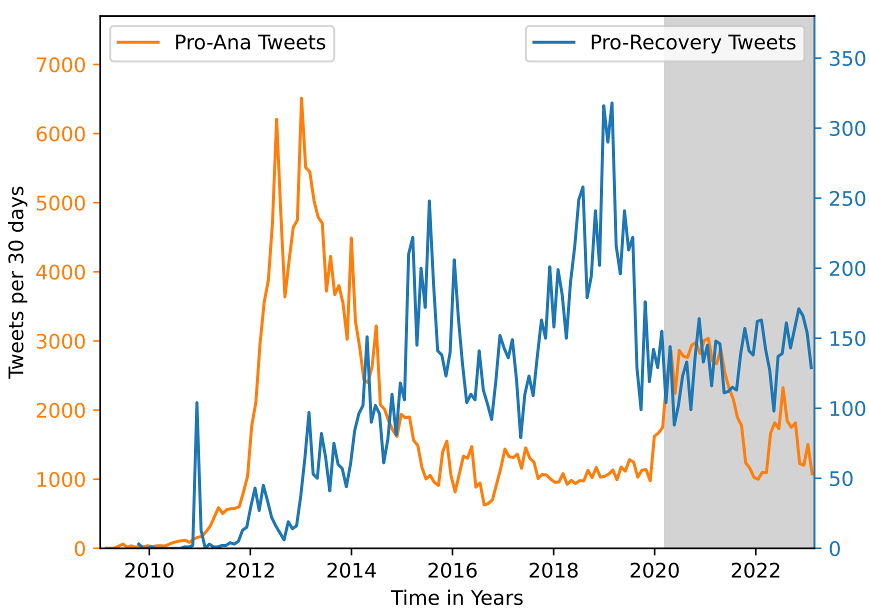
Descriptives Metadata (RQ1)

Tweet Volume

Figure 3 shows the relative frequency and volume of tweets allocated to the PA and PR communities over time. The PA community showed an overall greater volume of tweets than the PR community. Looking at their respective peaks in tweet volume, the number of PA tweets posted peaked in the beginning of 2013, while PR peaked later in mid 2019. When considering the relatively long timeframe, both peaks are likely to represent the height of an overall trend of the communities' popularity. Taking the tweet content itself in to account, suggested that the exact peak moments are not simply *random*, but both occurred during, and are thematically concerned with the Christmas holidays – a time that is especially difficult for individuals with anorexia due to frequent and expected eating. Interestingly, the amount of tweets in the PA community appeared to increase strongly around the start of the COVID-19 pandemic. The volume of tweets in the PR community remains roughly the same throughout the pandemic.

Figure 3

Frequency and Volume of Tweets over time (COVID-19 pandemic marked in grey).



Tweet Metrics

Table 2 shows the metrics of all tweets included in this study. The mean number of likes, retweets, quotes and replies were all higher for the tweets in the PA community. Most notably, an average PA tweet had almost twice as many likes and almost 3 times as many retweets. The standard deviation is also much larger for all tweet metrics, hinting at more outliers in PA tweets, which is also confirmed in the max values of all metrics. In total, these results suggest that PA members are more active in positive user feedback such as likes and retweets than PR members.

Table 2

Metadata metrics of all tweets in the PR and PA community.

	Pro-recovery (*k=17,839*)			**Pro-ana** (*k=288,773*)		
Tweet Metrics	*M*	*SD*	max	*M*	*SD*	max
Likes	2.98	30.73	3085	5.26	152.64	79806
Retweets	0.49	3.42	325	1.40	64.34	34344
Quotes	0.05	0.45	42	0.06	6.45	2919
Replies	0.31	1.81	100	0.32	9.02	4714

User Metrics

Table 3 presents the user metrics of all the people that posted in either community. The average user in the PR community has fewer followers, but follows more than twice as many other accounts, and also posted more tweets than the average PA community member. The average PR user is part of more of Twitter's internal content lists than the average PA users, suggesting lower levels of seclusion from other content/communities in the PR community.

Community Overlap

In terms of tweet overlap, it was found that a total of 259 tweets occurred in both communities. This amounts to only 0.08% of all tweets that were scraped from both communities. Further, it was also found that 457 users posted in both communities. This means that 11.64% of total PR users also posted tweets using a PA hashtag, while only 0.59% of total PA members also posted a tweet with a PR hashtag. As the total number of user overlap is greater than the tweet overlap, it can be concluded that most parts of user overlap occurs due to independent, unique tweets in either community.

Table 3

Metadata metrics of all users that posted within the PR and PA community.

	Pro-recovery (*k=17,839*)			**Pro-ana** (*k=288,773*)		
User Metrics	*M*	*SD*	max	*M*	*SD*	max
Followers	1980.38	11596.90	476192	2795.46	116969	19468000
Following	1045.58	5652.75	276292	441.05	1875.45	138386
Tweets	13197.25	40159.65	745368	10933.5	42392.8	3428500
Listed count	37.70	157.85	3669	20.11	581.21	106314

Extracted Sentiment (RQ2)

Table 4 shows the distribution of sentiment in all scraped tweets in the PA and PR communities. PA tweets were both more positive and negative, while PR tweets were more often neutral. This indicates that PA tweets tend to be higher in both positive and negative valence, while PR tweets are more neutral in their underlying sentiment.

Table 4

Sentiment distribution of pro-anorexia and pro-recovery tweets.

Community	**Positive**	**Neutral**	**Negative**	**Weighted average (-1.0 - 1.0)**
Pro-anorexia	44.5%	35.5%	20%	0.256
Pro-recovery	41.8%	43.9%	14.3%	0.273

Sentiment Over Time (RQ3)

Figure 4 shows the overall, weighted average, developments in sentiment of both communities over time. When relating these to the tweet frequencies, it can be seen that greater short-term variance in sentiment coincides with tweet frequency as the fewer tweets are posted, the stronger the sentiment fluctuates. When it comes to the overall development, the sentiment of PR tweets appears more unstable in short-term developments but more stable over the long-term than that of the PA community.

When looking at their development in relation to each other, more precise differences in sentiment between PA and PR can be observed. First, except for a brief period at the end of 2011, PA tweets mostly have a more positive sentiment until 2015. From then on, PA and PR have similar, positive, sentiments that slightly decrease until 2017. From 2017 onwards, the PA sentiment noticeably drops, turning negatively until the start of the COVID-19 pandemic. Interestingly, the PA sentiment starts rising again with the beginning of the pandemic, reaching its peak in 2022. The overall sentiment of PR tweets slightly but continuously decreases throughout the pandemic.

Figure 4

Sentiment of Tweets over time (COVID-19 pandemic marked in grey).

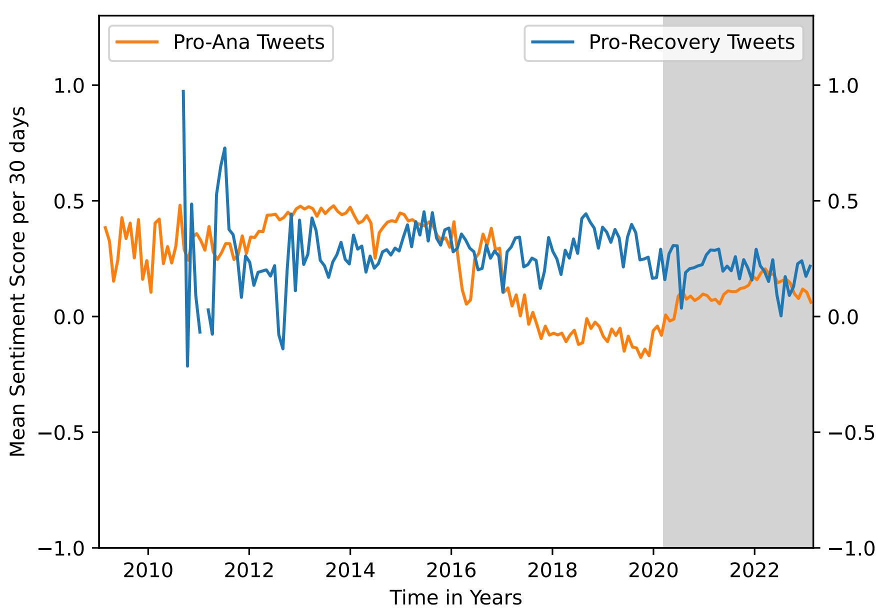


Figure 5 offers an even more detailed look into the development of each respective sentiment category within the PA and PR communities over the weighted average sentiment. Figure 5a and figure 5c show the total number of tweets with each sentiment that were posted over time. Figure 5b and figure 5d show the development of the likelihood for each tweet to have a certain sentiment over time.

Looking at the PR tweets in greater detail (Figure 5a & 5b), the proportion between sentiments is mostly stable. While some short-term changes can be noted, there is no apparent reason for these developments when investigating the content and metadata of the tweets, indicating a more broad shift in sentiment that does not seem to be influenced by one specific factor. From the onset of the COVID-19 pandemic, a general trend could be seen where positive and negative sentiments decreased while neutral sentiment increased.

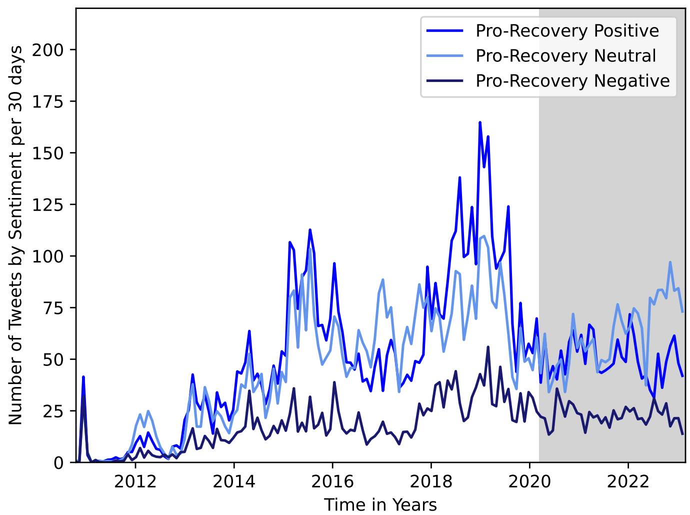
For the PA tweets (Figure 5c & 5d), the sentiment developed more stably with larger, overall trends. The change of PA sentiment from positive to negative appears to be caused by an increased activity of a select few users (likely bot accounts) repeatedly posting the same or similarly structured tweets, both on `#thinspiration` and on motivating users to vote in a community poll on why people decide to lose weight. With the beginning of the COVID-19 pandemic, PA sentiment became significantly more neutral, slightly more positive and less negative.

When comparing the developments of sentiment over time, a connection between the popularity of a community and the sentiment of their tweets can be seen. PA has most positive overall sentiment at its peak of popularity in 2013 as well as at the start of COVID-19 pandemic where its popularity also increased. When ignoring the strong fluctuation caused due to a lack of tweets at the beginning of PR, it shows the highest sentiment at its peak in popularity in 2019. Interestingly, the more popular PR is, the higher its sentiment, but also the lower the popularity and sentiment of PA tweets. In summary, this indicates a positive relationship of sentiment and popularity within, and a negative relationship between the sentiment and popularity between the communities.

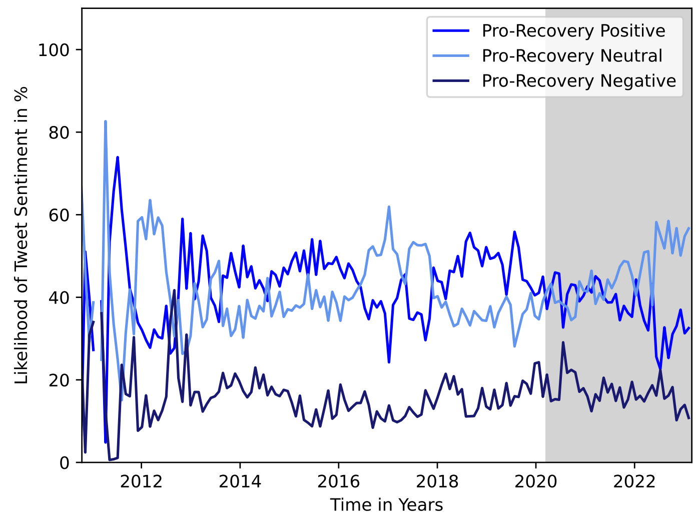
Figure 5

Number (a, c), and likelihood (b, d) of tweet sentiments in pro-ana and pro-recovery tweets.

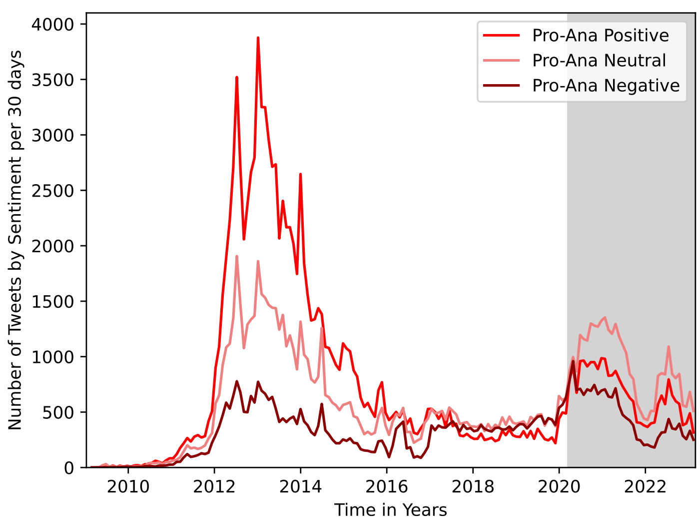
(a) *Volume of PR Tweets by Sentiment*



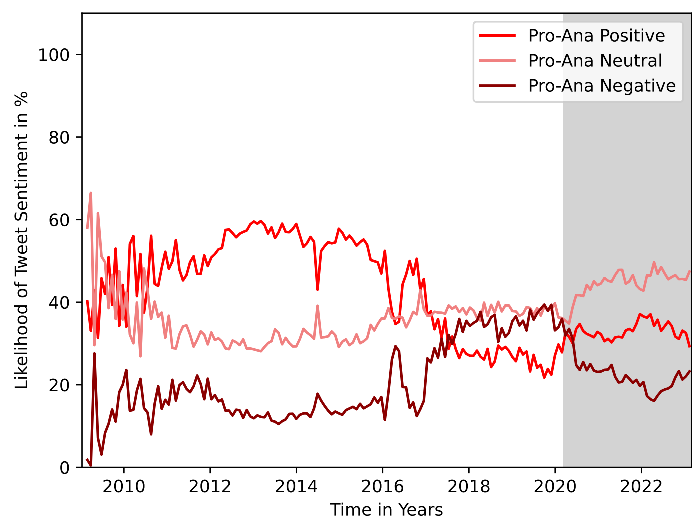
(b) *Likelihood of PR Sentiments*



(c) *Volume of PA Tweets by Sentiment*



(d) *Likelihood of PA Sentiments*



Note. Begin of the COVID-19 pandemic marked in grey.

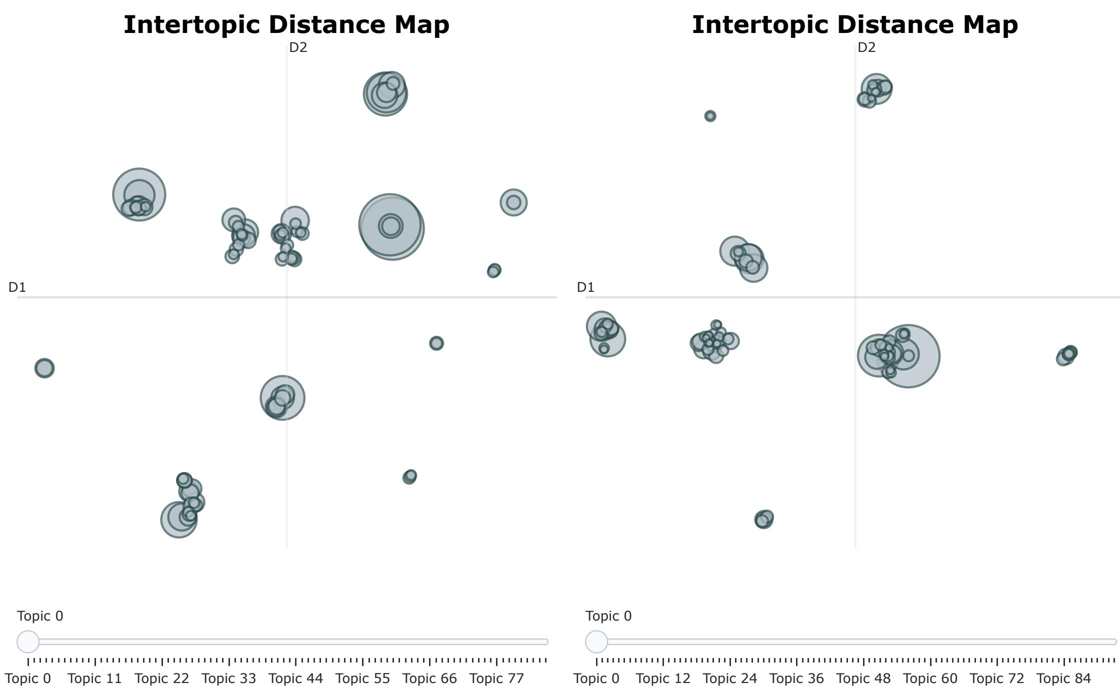
Extracted Topics (RQ4)

A full overview of all topics, their frequency and topic terms can be accessed through this link for PA, and this link for PR, or via Appendix A. The hierarchical relationship of all topics can be found in Appendix B1 & B2. Each figure in this section is designed to be interactively engaged with. The link to the interactive online versions can be accessed by clicking on the figure of interest or through Appendix A.

Figure 6

Intertopic Distance map for pro-recovery (a) and pro-anorexia (b) tweets.

(a) *Pro-recovery* (b) *Pro-anorexia*



Note. For the interactive versions of the figures click on the respective image or visit Appendix A.

Intertopic Distance Pro-recovery & Pro-anorexia

Figure 6a shows all topics as represented in a two-dimensional intertopic distance map. Each dot represents a topic, the larger the dot, the more tweets are associated with this topic. The closer the dots, the more similar they are to each other. Clusters of dots can, thus, be seen as overarching themes that connect multiple topics.

Intertopic Distance Pro-recovery. In total, the final topic representation contained 85 topics that cluster into 12 overarching themes. The largest topic was found in 906 tweets, while the smallest included topic is found in 21 tweets. Most clusters consist of multiple, similarly-sized topics. Four clusters contain only a few, small topics.

Intertopic Distance Pro-anorexia. In total, the final topic representation contains 93 topics that cluster into 7 overarching themes. The largest topic can be found in 24.999 tweets, the smallest topic in 320 tweets. In comparison to PR there are, thus, fewer themes that contain more topics. Furthermore, there are a few larger and many more mid-to-small-sized topics.

Topic Overview - General Observations

PR and PA tweets appeared to differ in the focus of their topics. In PR tweets, a topic is something that is actively discussed on a construct level (e.g. discussing what love is). In PA, topics exist more on a narrow emotion or subject level (e.g. “I love this”). Further, topic labels and the terms used in them closely resemble their original meaning in the topics discussed in PR tweets. For the PA topics, however, many concepts are isolated from their original meaning or are given new meaning in the context of anorexia. For example, “fashion” in the PA community mainly consists of *Victoria Secret* (a brand focussed on thin ideals and often strongly tied to anorexia and other EDs), but not actually of discussing fashion or other brands in general. “Transformation” always concerns weight loss and never a transformation outside of a weight context. Further, the combination of “normal” words with anorexia terms would create new meanings for these words. For example, the term *bedtime* was used in combination with *thinspiration* to refer to people “sending good night thinspiration” rather than wishing someone a good night. The PA topic also showed active seclusion and sub-communities where bedtime was restricted to *anasisters* so that it is only meant in the context of talking to *anasisters*. Interestingly, when looking at the similarity of topics, seemingly similar concepts like “Goals” (referring to general things to be achieved) and “Bodygoals” (referring to weight loss goals) mark two distinct concepts that are far from each other on a content level and show little engagement between the two groups.

Top 11 Topics Pro-recovery & Pro-ana

The 11 largest topics of the PR and PA community are shown in Table C1 and C2 respectively. It was chosen to present and compare these 11 topics as this amount enables a more detailed overview of the largest content within the two communities, while remaining interpretable in quantity. PR has almost 10% more outliers than PA. Furthermore, the size of PA topics is not just absolutely but also relatively greater, meaning that the top 11 PA topics contain more tweets than the top 11 PR topics.

When investigating these topics on a content level, it appears that none of the most prevalent PR topics centre around anorexia specifically. Instead, more general concepts, like recovery, EDs, or internet content appear most relevant. In contrast, all of the top 11 PA topics are either exclusively or at least heavily related to anorexia. Multiple topics revolve around different kinds of *thinspiration*, a term coined by the PA community. Further, heavy fixation on bodily features such as *bones*, or *weight loss* and *weight progress* also separate these topics much more from other EDs or more general communities. Another prominent feature of the most relevant PA topics is the association with others with anorexia in the topics *Anasister* and *Mutuals*.

Term-ranks. Ultimately, each topic is made up of different words or so-called terms that represent this topic. The relevance of each term differs per topic. Investigating the relevance of the terms per topic was not only useful for labelling the topics as they were, but also indicates how broadly spread a topic is in the terms that constitute it. Appendix B3a and B3b show the decline of term-relevance for all topics. When comparing the two, it can be seen that the first topic terms of PR topics are mostly less relevant than the first terms of PA. However, the PR terms decline less than those of PA. This indicates that the first one to three terms most strongly define a PA topic, while about five terms of similar relevance explain mostly a PR topic.

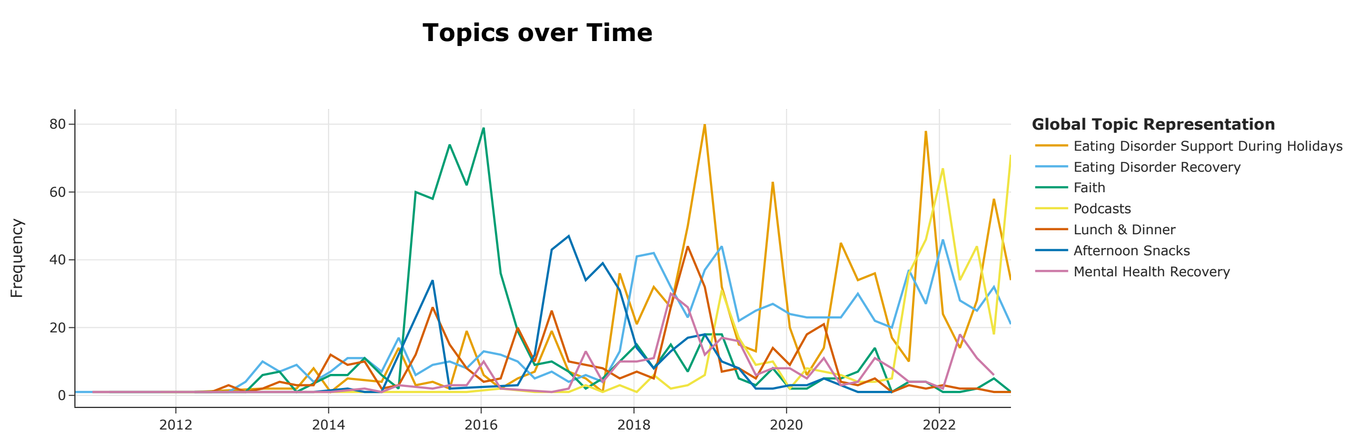
Barchart Topic Term Relevance. Based on the finding that the first five terms mostly explain a given topic, the terms of the top topics were explored in greater detail. Figure B4 shows multiple bar charts detailing the top topics and the representativeness of the terms that constitute them. When comparing PR and PA the aforementioned, stronger, decrease in term relevance of PA topic terms is confirmed. Overall, PA topics are much more strongly determined by the first few terms, with variations of *thinspiration* and *anasister* being most extreme. Contrastingly, the PR topics are explained more equally by the first five terms, with *ED Support During Holidays* and *Vegan Fitness* showing almost no drop-off in term relevance.

Dynamic Topic Extractions - Topics over Time (RQ5)

Pro-recovery Topics over Time

Figure 7

Pro-recovery topic frequency over time.



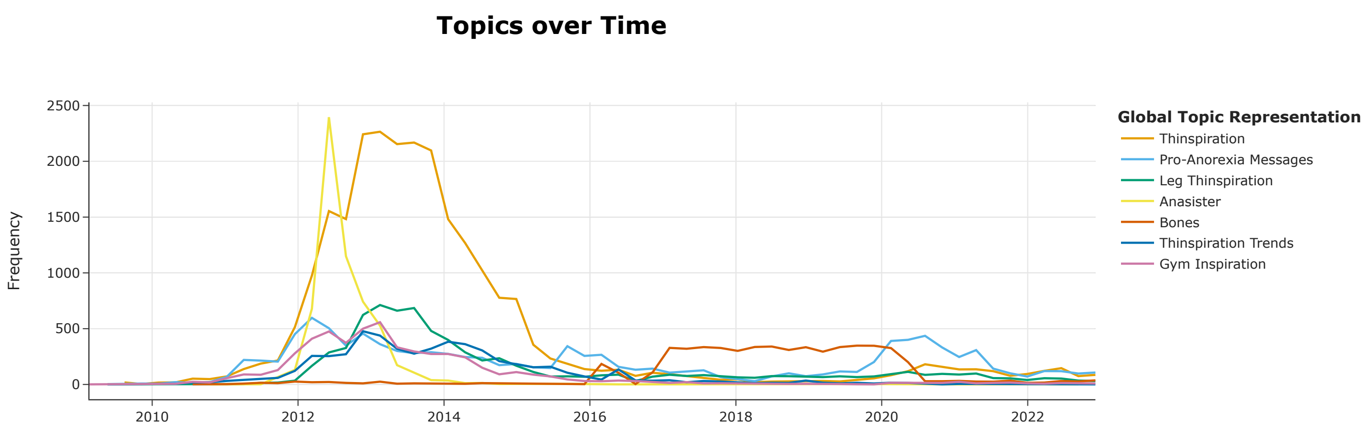
Note. For the interactive version, click on the image or visit Appendix A

Figure 7 shows the development of the seven largest topics by frequency over time. Generally, the topics heavily fluctuated in popularity, with few being more stable over time. The largest topic, *Eating Disorder Support During Holidays*, tended to periodically peak shortly before and during the Christmas holidays. *Eating Disorder Recovery* increased in relevancy in 2018 and stayed roughly similarly relevant. Another interesting finding is the strong peak of *Faith* from 2015 to 2016 and the peak of afternoon snack in 2015 and 2017, after which both topics largely decreased in relevancy. Perhaps the most recently relevant topic *Podcasts* strongly increased in the second half of 2021.

Pro-anorexia Topics over Time

Figure 8

Pro-anorexia topic frequency over time.



Note. For the interactive version, click on the image or visit Appendix A

Figure 8 shows the development of the seven largest PA topics over time. In comparison, the PA topics appeared to be more stable in their developments over time. While the overall peak of topic frequency in roughly 2013 aligns with the general peak in tweet frequency, two topics show a disproportionally large peak. For one, *Anasister* shows a singular peak in the middle of 2012. Further, *Thinspiration* shows an initial increase at the same time, with a peak throughout the whole year of 2013. After this, all topics decline in frequency. The topics *Bones* becomes relevant in 2017 and plateaus as the most frequent topic until 2020. After that, *PA Messages* took over for roughly the duration of the first COVID-19 waves. Most recently, of the top seven topics, all are found similarly frequently.

Discussion

This research is the first to use publicly available metadata and novel transformer-based text mining approaches to investigate pro-anorexia and pro-recovery communities on Twitter. In terms of differences in metadata between the communities (RQ1), it was found that the PA community shows a lot more overall tweets and user engagement through like, retweets and replies. The PR community shows fewer tweets with less user engagement and a broader range of engaged content. For the differences in overall sentiment (RQ2), PA tweets were found to be less neutral and instead both more positive and negative in sentiment, while PR tweets were more neutral. Sentiment developments over time (RQ3) were more stable in the short-term but more variable in the long-term for PA tweets, while it was the other way around for PR tweets. Further, a relationship between the communities' sentiment and popularity was found where a higher sentiment in one community also coincides with a higher volume of tweets posted within this community, but also a more negative sentiment and lower tweet volume in the other community. The discussed topics in both communities differ (RQ4) in that the PA tweet topics more closely revolve around anorexia itself on an emotional/subject level and are defined by fewer concepts. PR tweet topics were more on a construct level and focused more on general topics outside of AN while being defined by more concepts. Topic developments over time (RQ5) show the PA community to shift from initial strong peaks in relevance of the topics *Anasister* and *Thinspiration* to most recently no clear most relevant topic. Of the PR topics, *Eating Disorder Support During Holidays* annualy peaked in relevance and overall topic relevance shifted from *Faith* to *Afternoon Snacks* and most recently *Podcasts*.

Interpreting the Findings

Metadata Findings

Investigating the metadata, the first apparent difference was that more than 16 times as many tweets with PA hashtags than tweets with PR hashtags were found after excluding 77.252 (80%) PR tweets that were not anorexia specific. This may indicate that

PR is not as focussed on anorexia as the PA community, but instead operates as a more general mental health movement, targeting a broader audience of EDs, or even general mental health recovery. This would also be in line with the finding that PR users were part of more content lists, indicating a greater range of engaged content. Prior research by authors like Sukunesan et al. (2021) also supports this notion, regarding PR as a general recovery community for all EDs. PA tweeters, on the other hand, had a smaller listed count, underlining that they may be engaging with less content outside of anorexia. This confirms previous findings by authors like Branley and Covey (2017) that stated anorexic weight loss to be a central tenet of the community. This could mean that for PR to be more attractive for people that currently engage themselves with PA content, it may need a stronger focus on anorexia itself.

The PR community was found to be lower in their mean number of likes, retweets, quotes and replies, while having more tweets and a lower followers-to-following-ratio. Considering classifications like that of Daouadi et al. (2018), these metrics may hint towards the PR community being more “institutionalized”, meaning more driven by accounts led by organisations, such as healthcare institutions or awareness campaigns for example. PA, on the other hand, shows more intensive engagement in their relation of likes, retweets, replies and lower listed count, suggesting being more driven by individual, community oriented accounts (Daouadi et al., 2018). This finding matches previous research that postulated that PA members focus strongly on social adherence and interaction (Lai et al., 2021). To form helping relationships with anorexic individuals, an actor needs to convey “empathy; positive regard and acceptance; warmth; commitment; trust; genuineness; and be non-judgemental” (George, 1997, p. 902). It may be questionable whether organisations can convey these qualities in the same way that individual interaction can.

The increase of anorexia-related hospitalisations, described in literature, appeared to coincide with an increase in PA tweet volume throughout the first COVID-19 wave in this study (Matthews et al., 2021). As PR tweets did not increase in frequency, this may indicate that the PR community could be less relevant for individuals with acute AN.

Sentiment Findings

Overall Sentiment. The finding that PA tweets tended to be less neutral and both more positive and negative than PR tweets may be related to pro-anorexia being a polarising topic where little neutral content exists. For one, PA tweeters may speak positively about PA content, while outside actors speak negatively of it, in turn creating more polarising Twitter messages. However, it is also important to consider that the extreme lies at the very heart of PA in that its universal aim is to promote extreme thinness and condemn weight gain. PA members may, thus, present positive sentiments towards weight loss, but negative sentiment towards weight gain. Furthermore, PA members actively use and expect polarisation, for instance to trigger individuals into anorexic behaviour (*meanspiration* vs *sweetspiration*), which has also been confirmed to skew the sentiment in other research studies (Achilles et al., 2022). In between these extremes, there is little room for neutral sentiment to emerge. Bell (2007) also explains that through a combination of in-group dynamics as well as visual anonymity, online communities such as PA are inclined to polarise in their sentiment. Pre-existent tendencies may, thus, have increased through online interactions, leading to greater polarisation. The fact that PR tweets were more frequently neutral in sentiment may relate to its greater focus on providing factual information and increasing awareness, as has been found by Branley and Covey (2017). Further, the aforementioned finding of greater institutionalisation arguably leads to a more nuanced and neutral language being used in the PR community.

Sentiment Over Time. The stronger short-term variability of PR tweet sentiment over time is most likely best explained by the lower tweet volume in the PR community. However, global long-term trends should still be apparent, suggesting that, in contrast to the PA community, there is less change to PR sentiment. It was also found that the more tweets were posted in one community, the higher its positive sentiment and the lower its negative sentiment, but inversely the lower the volume of tweets in the other community and the lower the positive sentiment and the higher the negative sentiment of these PR tweets. This relationship between tweet frequency and sentiment may either imply that both sentiment and tweet volume are influenced by the same outside factor, or that one influences the other. However, when comparing tweet sentiment developments from 2017 to 2020, for example, a decrease in tweet volume can be seen to precede further decreases in tweet sentiment. The same observation can be made for the PR developments throughout the COVID-19 pandemic. Thus, it seems most apparent that sentiment developments follow the tweet frequency developments of a community. While the exact nature of this relationship is not fully known, it may potentially offer possible mechanisms of increasing the impact of PR communities by increasing its sentiment and their frequency of tweets.

Topic Findings

Overall Topics. The overall clustering of the themes and general focus of topic content strongly support the previously presented metadata finding that PA tweets are almost entirely concerned with anorexia and subsequent weight loss, while PR concerns itself with a broader range of content that is more independent of anorexia. The greater number of topic outliers further postulates a broader range of discussed topics in comparison to PA, providing even more evidence for this finding. This also fits the aforementioned indication that PR may be less attractive for individuals that suffer from acute AN. By having less direct focus on anorexia, PR may be less attractive for anorexic individuals to engage with on a content level as well. Further, the isolation and redefinition of concepts in the PA community indicate that the seclusion of PA does not only manifest on a content level but also on a linguistic and contextual level. Not only do PA community members focus strongly on discussing anorexia, but they utilise their own terms and concepts that are not immediately apparent from outside the community. This finding is in line with previous research on PA that described the community having a distinct lingo that people outside the community have no understanding of (Guthrie, 2013).

Topics Over Time. Over time, PR appeared to become more practically oriented, with greater focus on *Eating Disorder Recovery* or, most recently, recovery content via *Podcasts*. However, this shift did not coincide with increases in tweet volume. When integrating the PR topics over time with the tweet volume, it appears that *Eating Disorder Support During Holidays* directly correlates with the number of PR tweets. Thus, providing recovery support during the holidays arguably constitutes a relevant aspect of PR communities on Twitter. Christmas holidays especially can be difficult for individuals with anorexia, as they may relapse into stronger anorexic behaviours through being repeatedly confronted with food and eating customs (Dannibale, 2014). Considering that this risk regards relapse of individuals that have already overcome an acute phase, further strengthens the previous notion that, other than PA, PR more strongly manifests in providing assistance for individuals not currently suffering from acute anorexia.

When relating PA topics over time with tweet frequency, PA tweets appear to be more frequently posted when they also follows clear and distinct trends (e.g. *Thinspiration/Anasisters*, or later *Bones*). When there is no clear topic trend, PA also had fewer tweets. Group cohesion in terms of discussed topics, thus, seems to be an important factor for PA tweet volume. With no current main topic in the PA community, group cohesion and its subsequent tweet volume have noticeably decreased. As this has been a trend for more than a year, it can be assumed that this is a longer lasting change, rather than a short-term fluctuation. A plausible explanation for the finding that the PA community lost cohesion and had a lower number of tweets, even though COVID-19 was found to have led to significant increases in anorexia-related hospitalisation and ED-related SM content, is that PA members converted to other SM sites (Jordan et al., 2021; Matthews et al., 2021; Pruccoli et al., 2022). More specifically, *TikTok*, a short-form-video-based SM site has not only become the most used app by teenagers in general, but was also found to be strongly associated with PA community content (Business Wire, 2023; Jordan et al., 2021; Pruccoli et al., 2022).

Strengths, Limitations & Future Research

This research heavily utilised computer-driven analyses in the form of artificial intelligence to replace many steps of human labour in analysing tweets. This had the advantage of allowing analyses on a much larger scale, painting a more complete and objective image. Further, the methods used are more replicable than human-driven qualitative analyses. This allowed insights that could not otherwise be generated, constituting a major strength of this research. It may, however, also be argued that as these tools are trained through unsupervised learning on incomprehensibly large datasets, they may contain certain biases that are inconceivable to humans. This way, BERTopic or roBERTa-based-sentiment may have introduced some level of subjectivity or bias to the results. This "black box" nature of AI is an important factor, and it was considered in the extent of inferences that were made based on the results.

Another strength of the utilised transformers is their nearly endless flexibility in tweaking them for the best results. Almost any parameter in their pipeline can be changed, influencing the results of the analyses. However, this also introduces the difficulty of finding a balance between effectiveness vs efficiency, complexity vs simplicity, automation with AI vs reproducibility. Every decision has its advantages, but also introduces drawbacks. The fact that clear topic trends were identified both stably and over time, and even sarcastic or ironic tweets could be analysed accurately on their sentiment, suggests that this study succeeded at finding the right balance for the given data and application. However, as the utilized transformer models are relatively new, future research may shed light on how the models' parameters could be further tweaked to generate the best results.

In this study, 17 years' worth of Tweets in both the PR and PA communities were investigated. Throughout this time, many changes in language, discussed topics and sentiment were found. It could be argued that, as these were past developments, they are not essential to our understanding of current PA communities and, that with such a large scope of this study, nuanced details of its contemporary manifestation are lost as a consequence. However, when investigating recent tweets in the dataset of this study, many users can be seen referencing old tweets and past events. Dated content, partly from SM sites that have been mostly abandoned by contemporary PA communities, are frequently re-used and seldom regarded as “outdated”. Furthermore, by investigating the full history of PA communities on Twitter, a unique perspective on its developments and trends could be generated. This perspective does not only benefit our understanding of current PA communities, but may also help us to understand future developments. Investigating this long timeframe with special focus on developments over time can, thus, be considered a strength of this research, rather than a limitation. In addition, dynamic topic modelling specifically marks a very novel approach to topic modelling, and this study is one of the first in the field of psychology to adopt and apply this approach.

This study used text-mining methods for understanding anorexia communities, putting the focus on the text of tweets while disregarding any other content, like pictures or videos attached to tweets. While this yielded promising results, the sharing of images is also considered to be an important characteristic of PA communities (Lerman et al., 2023). PA users post pictures to share weight-loss progress, diet/workout plans, or distribute inspiration content like thinspiration (Borzekowski et al., 2010; Chancellor, Mitra, et al., 2016; Lerman et al., 2023). Future research may make focus on multi-modal data, such as images, to potentially further inform text-mining findings. In this end, promising tools are already in active development. The developer of BERTopic combined its features with another AI that has been trained on recognizing and categorizing images into a tool called “Concept” ↗. This is especially promising as image content is fairly repetitive with focus on specific themes in the PA community. Furthermore, future developments may also allow the application of similar tools on video content to extend the use of AI to short-form video platforms like *TikTok*.

The major strength of this research lies in its approach to open science. All of the Python code ↗ that was specifically written to conduct the analyses is documented and made publicly available to be used by anyone (Appendix A). This has the advantage of lowering the barrier of entry for every future research study aiming to replicate or repurpose the methods used in this study. Furthermore, all of the datasets from this study can also be openly accessed and used (Appendix A). As these constitute data on all tweets that were ever posted using the hashtags mentioned in Table 1, it also serves as an archive of the PA and PR communities on Twitter from its founding until the 1st of February 2023. The results of this research exceed simply answering the research questions, but also constitute a foundation for future research on PA and PR research to build on.

Implications

As this research was of exploratory nature, its implications are on a more theoretical than practical level. For one, *BERTopic* and *roBERTa-base-sentiment* have been shown to novel insights for Twitter data. Their flexible nature has been demonstrated to allow for effective parameter tuning for more accurate results. However, this fine-tuning process constitutes a balancing act between increasing the quality, depth, and range of insights, while also keeping analyses replicable, understandable and resource efficient (e.g. in terms of time and computing needed). It has also been demonstrated how developments in sentiment and topics over time can be investigated and how they may facilitate contemporary understanding. Future research can be based on the design of this study, as well as the openly accessible python code and datasets. As the applicability of these analyses has been demonstrated, future research may also build upon the acquired insights and data to more elaborately investigate single aspects, such as topics over time etc., in greater detail.

Conclusion

This research investigated how PA and PR communities manifest and contrast on Twitter. To this end, metadata and newly emerging transformer-based text mining applications were used to explore interactional dynamics, tweet sentiment, and discussed topics, both globally and over time. Overall, clear differences between the two communities were found. The PA community appeared to display more social interaction and conversations of anorexia-related topics. PA tweeters tend to be more secluded from other communities, using a lingo that is unique to this community and heavily bound to anorexia and weight loss. PA tweets also tended to be more extreme in sentiment, making use of polarising language and topics. The PR community on the other hand appears more institutionalised, discussing more topics outside of anorexia in a more neutral language. This may make it appear less genuine as a community and less interesting to individuals in an anorexic crisis. While PA is much more frequently posted in as a result, there is a relationship in both communities where more positive language is related to increased tweet volume in one community but more negative language and lower tweet volume of the other community.

To promote the transition from harmful PA to more beneficial communities, it remains important to expand our understanding of them. While this study provided novel insights into the PA and PR community, future research is needed to further deepen these insights as well as test potential methods of intervention based on them. Computer-driven analyses like those used in this research offer an important advantage that allows utilizing the large quantity of SM data to the advantage of research and beneficial interventions alike.

References

Achilles, L., Mandl, T., & Womser-Hacker, C. (2022). “Meanspo Please, I Want to Lose Weight”: A Characterization Study of Meanspiration Content on Tumblr Based on Images and Texts. In A. Barrón-Cedeño, G. Da San Martino, M. Degli Esposti, F. Sebastiani, C. Macdonald, G. Pasi, A. Hanbury, M. Potthast, G. Faggioli, & N. Ferro (Eds.), *Experimental IR Meets Multilinguality, Multimodality, and Interaction* (pp. 3–17). Springer International Publishing. https://doi.org/10.1007/978-3-031-13643-6_1

Akundi, A., Tseng, B., Wu, J., Smith, E., Subbalakshmi, M., & Aguirre, F. (2018). Text Mining to Understand the Influence of Social Media Applications on Smartphone Supply Chain. *Procedia Computer Science*, *140*, 87–94. https://doi.org/10.1016/j.procs.2018.10.296

Albalawi, R., Yeap, T. H., & Benyoucef, M. (2020). Using Topic Modeling Methods for Short-Text Data: A Comparative Analysis. *Frontiers in Artificial Intelligence*, *3*. https://doi.org/10.3389/frai.2020.00042

Arcelus, J., Mitchell, A. J., Wales, J., & Nielsen, S. (2011). Mortality Rates in Patients With Anorexia Nervosa and Other Eating Disorders: A Meta-analysis of 36 Studies. *Archives of General Psychiatry*, *68*(7), 724–731. https://doi.org/10.1001/archgenpsychiatry.2011.74

Arseniev-Koehler, A., Lee, H., McCormick, T., & Moreno, M. A. (2016). #Proana: Pro-Eating Disorder Socialization on Twitter. *Journal of Adolescent Health*, *58*(6), 659–664. https://doi.org/10.1016/j.jadohealth.2016.02.012

Assemblée nationale. (n.d.). N° 3481 - Proposition de loi de M. François Vannson tendant à interdire les sites Internet valorisant et faisant l’apologie de l’anorexie - XIIe législature - Assemblée nationale.

BBC News. (2012). Tumblr to ban self-harm and eating disorder blogs. *BBC News*.

Bell, V. (2007). Online information, extreme communities and internet therapy: Is the internet good for our mental health? *Journal of Mental Health*, *16*(4), 445–457. https://doi.org/10.1080/09638230701482378

Borzekowski, D. L. G., Schenk, S., Wilson, J. L., & Peebles, R. (2010). E-Ana and e-Mia: A Content Analysis of Pro–Eating Disorder Web Sites. *American Journal of Public Health*, *100*(8), 1526–1534. https://doi.org/10.2105/AJPH.2009.172700

Branley, D. B., & Covey, J. (2017). Pro-ana versus Pro-recovery: A Content Analytic Comparison of Social Media Users' Communication about Eating Disorders on Twitter and Tumblr. *Frontiers in Psychology*, *8*.

Business Wire. (2023). Piper Sandler Completes 45th Semi-Annual Generation Z Survey of 5,690 U.S. Teens.

Cachelin, F. M., & Striegel-Moore, R. H. (2006). Help seeking and barriers to treatment in a community sample of Mexican American and European American women with eating disorders. *The International Journal of Eating Disorders*, *39*(2), 154–161. https://doi.org/10.1002/eat.20213

Camacho-Collados, J., Rezaee, K., Riahi, T., Ushio, A., Loureiro, D., Antypas, D., Boisson, J., Espinosa-Anke, L., Liu, F., Martínez-Cámara, E., Medina, G., Buhrmann, T., Neves, L., & Barbieri, F. (2022). TweetNLP: Cutting-Edge Natural Language Processing for Social Media. https://doi.org/10.48550/arXiv.2206.14774

Casilli, A. A., Pailler, F., & Tubaro, P. (2013). Online networks of eating-disorder websites: Why censoring pro-ana might be a bad idea. *Perspectives in Public Health*, *133*(2), 94–95. https://doi.org/10.1177/1757913913475756

Chancellor, S., Mitra, T., & De Choudhury, M. (2016). Recovery Amid Pro-Anorexia: Analysis of Recovery in Social Media. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2111–2123. https://doi.org/10.1145/2858036.2858246

Chancellor, S., Pater, J. A., Clear, T., Gilbert, E., & De Choudhury, M. (2016). #Thyghgapp: Instagram Content Moderation and Lexical Variation in Pro-Eating Disorder Communities. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 1201–1213. https://doi.org/10.1145/2818048.2819963

Chowdhary, K. R. (2020). Natural Language Processing. In K. Chowdhary (Ed.), *Fundamentals of Artificial Intelligence* (pp. 603–649). Springer India. https://doi.org/10.1007/978-81-322-3972-7_19

Chung, A., Vieira, D., Donley, T., Tan, N., Jean-Louis, G., Gouley, K. K., & Seixas, A. (2021). Adolescent Peer Influence on Eating Behaviors via Social Media: Scoping Review. *Journal of Medical Internet Research*, *23*(6), e19697. https://doi.org/10.2196/19697

Cobb, G. (2017). "This is not pro-ana": Denial and disguise in pro-anorexia online spaces. *Fat Studies*, *6*(2), 189–205. https://doi.org/10.1080/21604851.2017.1244801

Dannibale, K. (2014). The Effects of the Holidays on Eating Disorders. *New Errands: The Undergraduate Journal of American Studies*. https://doi.org/10.18113/P8ne2159255

Daouadi, K. E., Rebaï, R. Z., & Amous, I. (2018). Organization vs. Individual: Twitter user classification. *LPKM*.

DCMS. (n.d.). Online Harms White Paper: Full government response to the consultation.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. https://doi.org/10.48550/arXiv.1810.04805

DeWitt, J., & Attia, E. (2017). Anorexia Nervosa. In T. Wade (Ed.), *Encyclopedia of Feeding and Eating Disorders* (pp. 13–16). Springer. https://doi.org/10.1007/978-981-287-104-6_51

Egger, R., & Yu, J. (2022). A Topic Modeling Comparison Between LDA, NMF, Top2Vec, and BERTopic to Demystify Twitter Posts. *Frontiers in Sociology*, *7*. https://doi.org/10.3389/fsoc.2022.886498

Elbagir, S., & Yang, J. (2019). Twitter Sentiment Analysis Using Natural Language Toolkit and VADER Sentiment. *Hong Kong*.

Ferguson, C. J., Muñoz, M. E., Garza, A., & Galindo, M. (2014). Concurrent and Prospective Analyses of Peer, Television and Social Media Influences on Body Dissatisfaction, Eating Disorder Symptoms and Life Satisfaction in Adolescent Girls. *Journal of Youth and Adolescence*, *43*(1), 1–14. https://doi.org/10.1007/s10964-012-9898-9

Fettach, Y., & Benhiba, L. (2019). Pro-Eating Disorders and Pro-Recovery Communities on Reddit: Text and Network Comparative Analyses. *Proceedings of the 21st International Conference on Information Integration and Web-based Applications & Services*, 277–286. https://doi.org/10.1145/3366030.3366058

Garman, A. N., Standish, M. P., & Kim, D. H. (2018). Enhancing efficiency, reliability, and rigor in competency model analysis using natural language processing. *The Journal of Competency-Based Education*, *3*(3), e01164. https://doi.org/10.1002/cbe2.1164

George, L. (1997). The psychological characteristics of patients suffering from anorexia nervosa and the nurse's role in creating a therapeutic relationship. *Journal of Advanced Nursing*, *26*(5), 899–908. https://doi.org/10.1046/j.1365-2648.1997.00421.x

Giordani, R. C. F., & Silva, F. S. (2021). The ethereal bodies of pro-Ana blogs: Emotional communities and spaces of sociability on the web. *Ciência & Saúde Coletiva*, *26*, 5293–5301. https://doi.org/10.1590/1413-812320212611.3.34522019

Glez-Peña, D., Lourenço, A., López-Fernández, H., Reboiro-Jato, M., & Fdez-Riverola, F. (2014). Web scraping technologies in an API world. *Briefings in Bioinformatics*, *15*(5), 788–797. https://doi.org/10.1093/bib/bbt026

Grabe, S., Ward, L. M., & Hyde, J. S. (2008). The role of the media in body image concerns among women: A meta-analysis of experimental and correlational studies. *Psychological Bulletin*, *134*, 460–476. https://doi.org/10.1037/0033-2909.134.3.460

Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure.

Guthrie, C. (2013). Social Support Found by Those Visiting Eating Disorder Websites May Lead Untreated Cases.

Hall, J. A., & Liu, D. (2022). Social media use, social displacement, and well-being. *Current Opinion in Psychology*, *46*, 101339. https://doi.org/10.1016/j.copsyc.2022.101339

Harmon, Roesslein, J., & other contributors. (n.d.). Tweepy. https://doi.org/10.5281/zenodo.7259945

Harper, K., Sperry, S., & Thompson, J. K. (2008). Viewership of pro-eating disorder websites: Association with body image and eating disturbances. *International Journal of Eating Disorders*, *41*(1), 92–95. https://doi.org/10.1002/eat.20408

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., Fernández del Río, J., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*, 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hassani, H., Beneki, C., Unger, S., Mazinani, M. T., & Yeganegi, M. R. (2020). Text Mining in Big Data Analytics. *Big Data and Cognitive Computing*, *4*(1), 1. https://doi.org/10.3390/bdcc4010001

Hosokawa, R., Kawabe, K., Nakachi, K., Soga, J., Horiuchi, F., & Ueno, S.-i. (2023). Effects of social media on body dissatisfaction in junior high school girls in Japan. *Eating Behaviors*, *48*, 101685. https://doi.org/10.1016/j.eatbeh.2022.101685

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Instagram. (n.d.). About Eating Disorders | Instagram Help Center.

Jonsson, E., & Stolee, J. (n.d.). An Evaluation of Topic Modelling Techniques for Twitter.

Jordan, G. L., García, M. D., Esteban, P. S., Martín, A. S., Sánchez, P. M., Cubo, M. J., Cano, M. M., Barrio, J. G. D., & Ayesa-Arriola, R. (2021). Tiktok, a vehicle for Pro-Ana and Pro-Mia content boosted by the COVID-19 pandemic. *European Psychiatry*, *64*(S1), S703–S703. https://doi.org/10.1192/j.eurpsy.2021.1862

Kherwa, P., & Bansal, P. (2018). Topic Modeling: A Comprehensive Review. *ICST Transactions on Scalable Information Systems*, *0*(0), 159623. https://doi.org/10.4108/eai.13-7-2018.159623

Lai, C., Pellicano, G. R., Iuliano, S., Ciacchella, C., Sambucini, D., Gennaro, A., & Salvatore, S. (2021). Why people join pro-Ana online communities? A psychological textual analysis of eating disorder blog posts. *Computers in Human Behavior*, *124*, 106922. https://doi.org/10.1016/j.chb.2021.106922

Lerman, K., Karnati, A., Zhou, S., Chen, S., Kumar, S., He, Z., Yau, J., & Horn, A. (2023). Radicalized by Thinness: Using a Model of Radicalization to Understand Pro-Anorexia Communities on Twitter. https://doi.org/10.48550/arXiv.2305.11316

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach.

Loureiro, D., Barbieri, F., Neves, L., Anke, L. E., & Camacho-Collados, J. (2022). TimeLMs: Diachronic Language Models from Twitter. https://doi.org/10.48550/arXiv.2202.03829

Matthews, A., Kramer, R. A., Peterson, C. M., & Mitan, L. (2021). Higher admission and rapid readmission rates among medically hospitalized youth with anorexia nervosa/atypical anorexia nervosa during COVID-19. *Eating Behaviors*, *43*, 101573. https://doi.org/10.1016/j.eatbeh.2021.101573

McInnes, L., Healy, J., Saul, N., & Grossberger, L. (2018). UMAP: Uniform Manifold Approximation and Projection. *The Journal of Open Source Software*, *3*(29), 861.

Mento, C., Silvestri, M. C., Muscatello, M. R. A., Rizzo, A., Celebre, L., Praticò, M., Zoccali, R. A., & Bruno, A. (2021). Psychological Impact of Pro-Anorexia and Pro-Eating Disorder Websites on Adolescent Females: A Systematic Review. *International Journal of Environmental Research and Public Health*, *18*(4), 2186. https://doi.org/10.3390/ijerph18042186

Park, E., Kim, W.-H., & Kim, S.-B. (2022). What Topics Do Members of the Eating Disorder Online Community Discuss and Empathize with? An Application of Big Data Analytics. *Healthcare*, *10*(5), 928. https://doi.org/10.3390/healthcare10050928

Pinterest. (n.d.). Community guidelines.

Pruccoli, J., De Rosa, M., Chiasso, L., Perrone, A., & Parmeggiani, A. (2022). The use of TikTok among children and adolescents with Eating Disorders: Experience in a third-level public Italian center during the SARS-CoV-2 pandemic. *Italian Journal of Pediatrics*, *48*(1), 138. https://doi.org/10.1186/s13052-022-01308-4

Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*.

Rich, E. (2006). Anorexic dis(connection): Managing anorexia as an illness and an identity. *Sociology of Health & Illness*, *28*(3), 284–305. https://doi.org/10.1111/j.1467-9566.2006.00493.x

Rouleau, C. R., & von Ranson, K. M. (2011). Potential risks of pro-eating disorder websites. *Clinical Psychology Review*, *31*(4), 525–531. https://doi.org/10.1016/j.cpr.2010.12.005

Schmidt, U., Startup, H., & Treasure, J. (2018). *A Cognitive Interpersonal Therapy Workbook for Treating Anorexia Nervosa: The Maudsley Model* (First). Routledge. https://doi.org/10.4324/9781315728483

Shelar, A., & Huang, C.-Y. (2018). Sentiment Analysis of Twitter Data. *2018 International Conference on Computational Science and Computational Intelligence (CSCI)*, 1301–1302. https://doi.org/10.1109/CSCI46756.2018.00252

Sidani, J. E., Shensa, A., Hoffman, B., Hanmer, J., & Primack, B. A. (2016). The Association between Social Media Use and Eating Concerns among US Young Adults. *Journal of the Academy of Nutrition and Dietetics*, *116*(9), 1465–1472. https://doi.org/10.1016/j.jand.2016.03.021

Silén, Y., & Keski-Rahkonen, A. (2022). Worldwide prevalence of DSM-5 eating disorders among young people. *Current Opinion in Psychiatry*, *35*(6), 362–371. https://doi.org/10.1097/YCO.0000000000000818

Statista. (n.d.). Global Twitter user age distribution 2021.

Sukunesan, S., Huynh, M., & Sharp, G. (2021). Examining the Pro-Eating Disorders Community on Twitter Via the Hashtag #proana: Statistical Modeling Approach. *JMIR Mental Health*, *8*(7), e24340. https://doi.org/10.2196/24340

Sullivan, P. F. (1995). Mortality in anorexia nervosa. *The American Journal of Psychiatry*, *152*(7), 1073–1074. https://doi.org/10.1176/ajp.152.7.1073

Swanson, S. A., Crow, S. J., Le Grange, D., Swendsen, J., & Merikangas, K. R. (2011). Prevalence and Correlates of Eating Disorders in Adolescents: Results From the National Comorbidity Survey Replication Adolescent Supplement. *Archives of General Psychiatry*, *68*(7), 714–723. https://doi.org/10.1001/archgenpsychiatry.2011.22

Tiggemann, M., Churches, O., Mitchell, L., & Brown, Z. (2018). Tweeting weight loss: A comparison of #thinspiration and #fitspiration communities on Twitter. *Body Image*, *25*, 133–138. https://doi.org/10.1016/j.bodyim.2018.03.002

van Hoeken, D., & Hoek, H. W. (2020). Review of the burden of eating disorders: Mortality, disability, costs, quality of life, and family burden. *Current Opinion in Psychiatry*, *33*(6), 521. https://doi.org/10.1097/YCO.0000000000000641

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., ... SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in python. *Nature Methods*, *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2

Wang, T., Brede, M., Ianni, A., & Mentzakis, E. (2017). Detecting and Characterizing Eating-Disorder Communities on Social Media. *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, 91–100. https://doi.org/10.1145/3018661.3018706

Wilson, J. L., Peebles, R., Hardy, K. K., & Litt, I. F. (2006). Surfing for Thinness: A Pilot Study of Pro–Eating Disorder Web Site Usage in Adolescents With Eating Disorders. *Pediatrics*, *118*(6), e1635–e1643. https://doi.org/10.1542/peds.2006-1133

Yadav, A., & Vishwakarma, D. K. (2020). Sentiment analysis using deep learning architectures: A review. *Artificial Intelligence Review*, *53*(6), 4335–4385. https://doi.org/10.1007/s10462-019-09794-5

Yom-Tov, E., Fernandez-Luque, L., Weber, I., & Crain, S. P. (2012). Pro-Anorexia and Pro-Recovery Photo Sharing: A Tale of Two Warring Tribes. *Journal of Medical Internet Research*, *14*(6), e2239. https://doi.org/10.2196/jmir.2239

Zhou, S., Zhao, Y., Bian, J., Haynos, A. F., & Zhang, R. (2020). Exploring Eating Disorder Topics on Twitter: Machine Learning Approach. *JMIR Medical Informatics*, *8*(10), e18273. https://doi.org/10.2196/18273

Appendix A

Hyperlinks

Python Code	https://juliusheinke.github.io/Master-Thesis/Notebook/
Datasets	https://github.com/juliusheinke/Master-Thesis/tree/main/Datasets

Pro-Anorexia Topics

All Topics	https://juliusheinke.github.io/Master-Thesis/pa/topics/
Topic Hierarchy	https://juliusheinke.github.io/Master-Thesis/pa/hierarchy/
Intertopic Distance Map	https://juliusheinke.github.io/Master-Thesis/pa/idm/
Term Ranks	https://juliusheinke.github.io/Master-Thesis/pa/term_rank/
Topic Term Relavence	https://juliusheinke.github.io/Master-Thesis/pa/barchart/
Topics over Time	https://juliusheinke.github.io/Master-Thesis/pa/tot/

Pro-Recovery Topics

All Topics	https://juliusheinke.github.io/Master-Thesis/pr/topics/
Topic Hierarchy	https://juliusheinke.github.io/Master-Thesis/pr/hierarchy/
Intertopic Distance Map	https://juliusheinke.github.io/Master-Thesis/pr/idm/
Term Ranks	https://juliusheinke.github.io/Master-Thesis/pr/term_rank/
Topic Term Relavence	https://juliusheinke.github.io/Master-Thesis/pr/barchart/
Topics over Time	https://juliusheinke.github.io/Master-Thesis/pr/tot/

Appendix B

Figures

Figure B1

Pro-recovery topic hierarchy (Click on figure or visit Appendix A for interactive version).

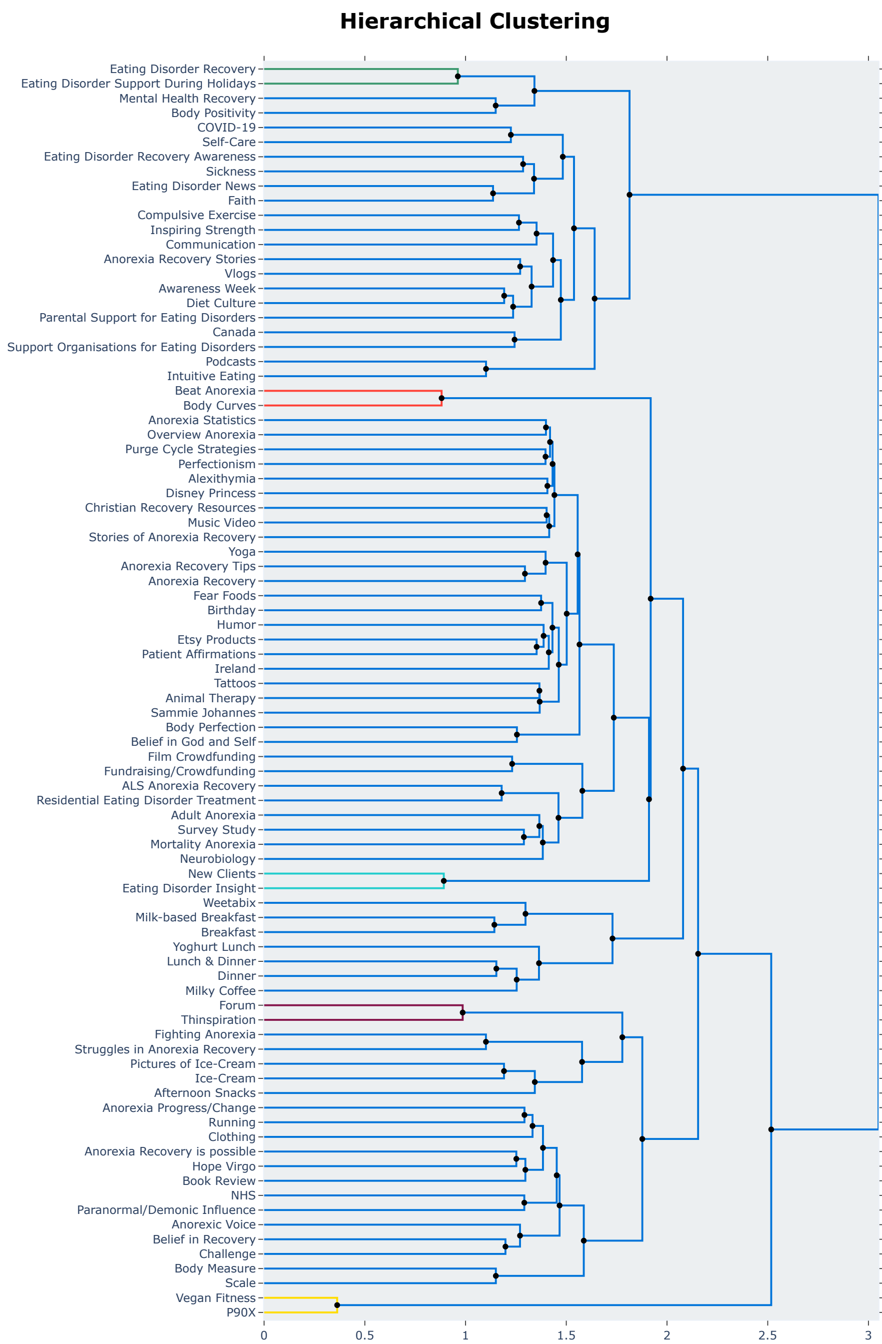


Figure B2

Pro-anorexia topic hierarchy (Click on figure or visit Appendix A for interactive version).

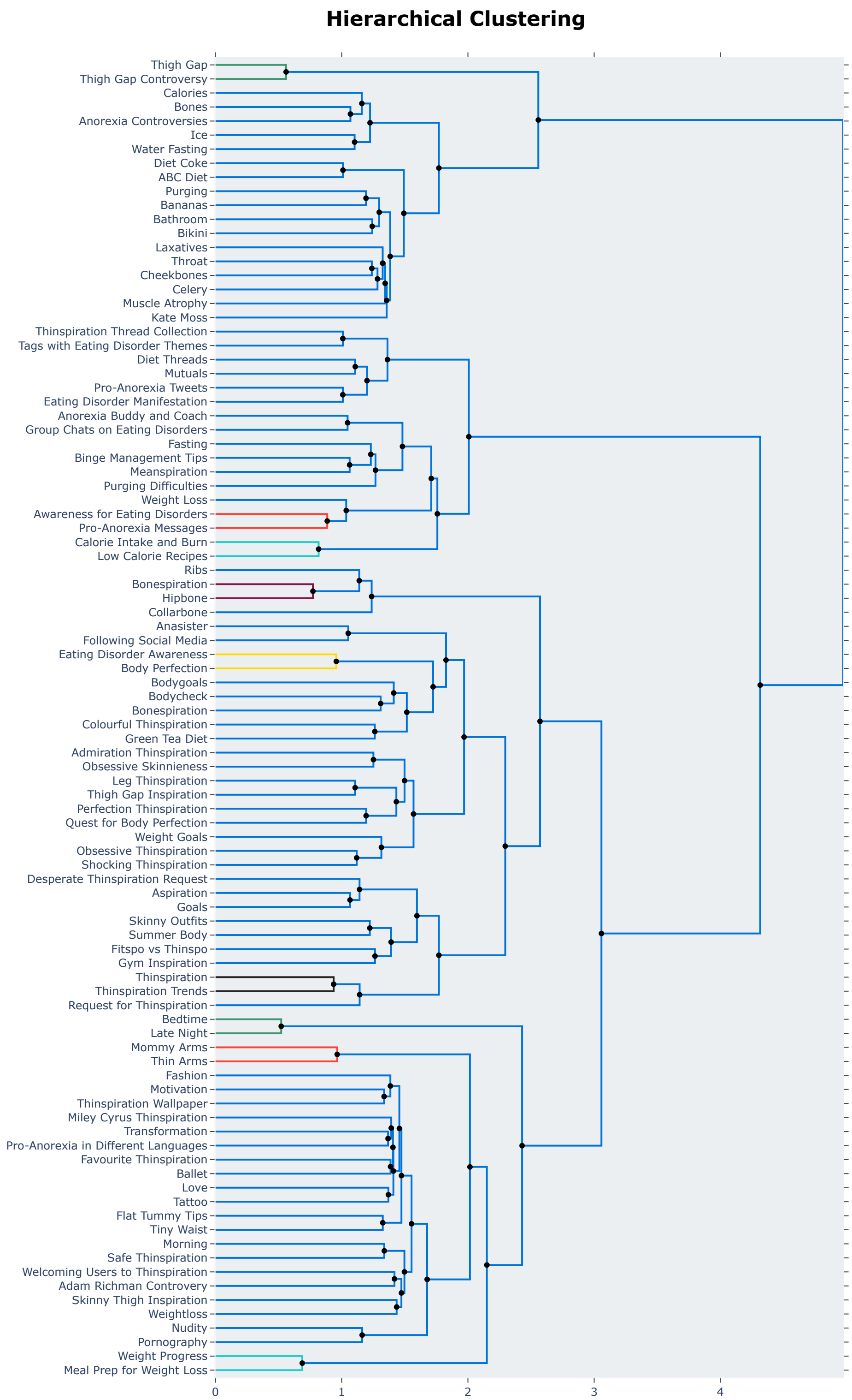
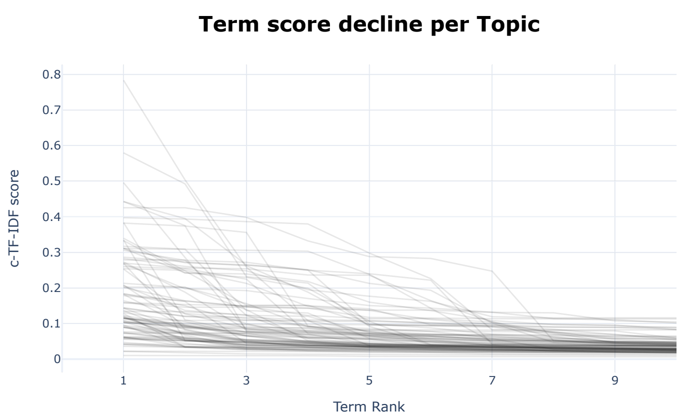


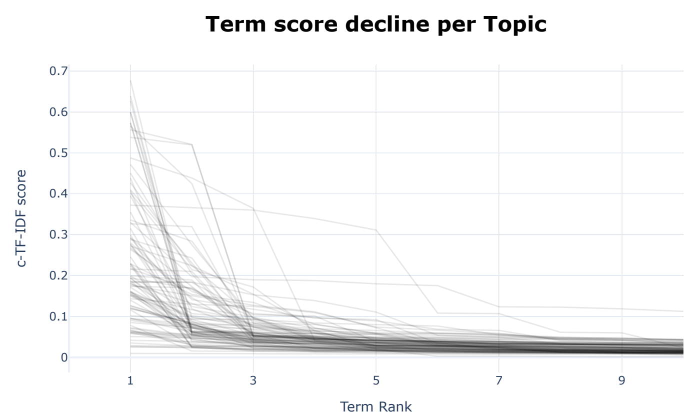
Figure B3

Term rank for all Pro-recovery (a) and Pro-anorexia (b) topics.

(a) *Pro-recovery*



(b) *Pro-anorexia*

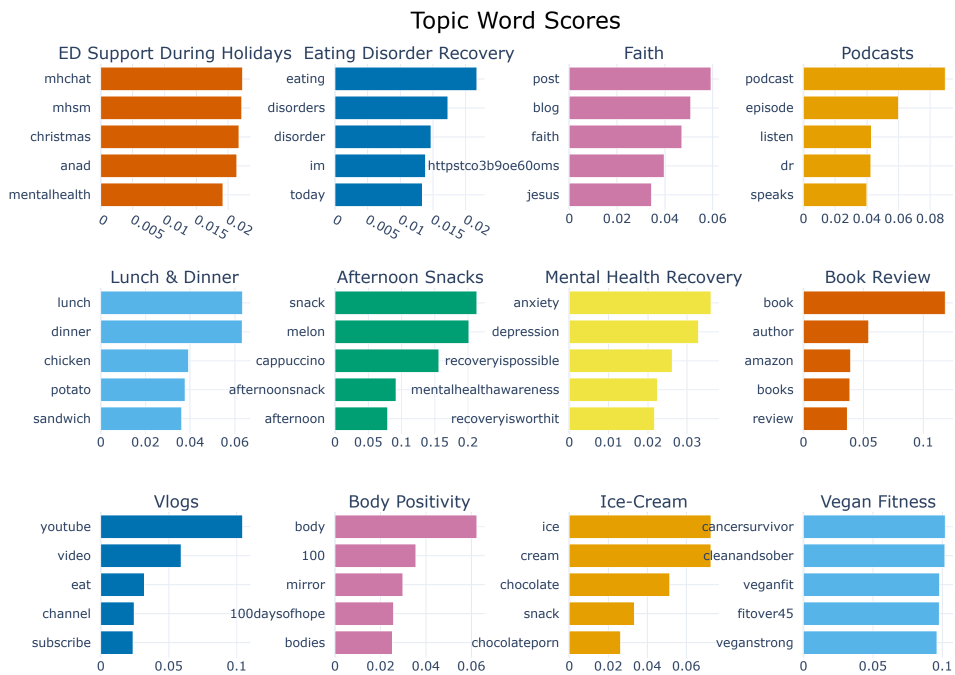


Note. For the interactive version, click on the respective image or visit Appendix A.

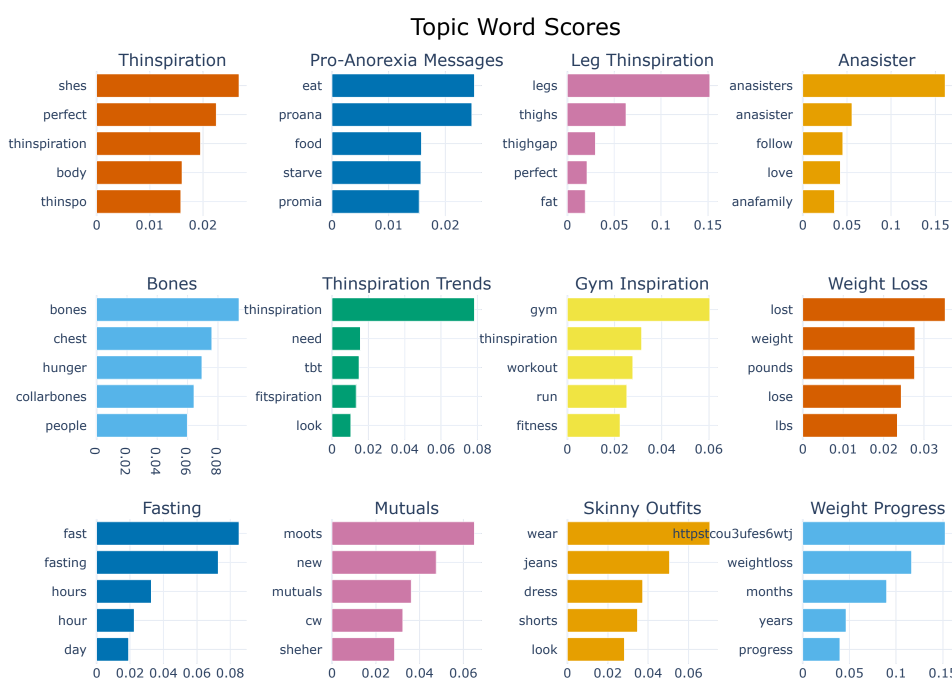
Figure B4

Barchart of Pro-recovery (a) and Pro-anorexia (b) Topic term relevance.

(a) *Pro-recovery*



(b) *Pro-anorexia*



Appendix C

Tables

Table C1

Top 11 most relevant Pro-recovery topics

Topic	Count	% of Total	Label	Topic Words
-1	9182	51.17	Outliers	anorexia| edrecovery| recovery| anorexiarecovery| bulimia| eatingdisorderrecovery| mentalhealth
0	906	5.08	Eating Disorder Support During Holidays	mhchat| mhsm| christmas| anad| mentalhealth| bingeeatingdisorder| ednos
1	876	4.91	Eating Disorder Recovery	eating| disorders| disorder| im| today| weight| ive
2	629	3.53	Faith	post| blog| faith| httpstco3b9oe60oms| jesus| new| god
3	445	2.49	Podcasts	podcast| episode| listen| dr| speaks| cristina| castagnini
4	434	2.43	Lunch & Dinner	lunch| dinner| chicken| potato| sandwich| peas| sweet
5	359	2.01	Afternoon Snacks	snack| melon| cappuccino| afternoonsnack| afternoon| anarecovery| fruit
6	291	1.63	Mental Health Recovery	anxiety| depression| recoveryispossible| mentalhealthawareness| recoveryisworthit| ptsd| bipolar
7	213	1.19	Book Review	book| author| amazon| books| review| reading| copy
8	175	0.98	Vlogs	youtube| video| eat| channel| subscribe| vlog| videos
9	170	0.95	Body Positivity	body| 100| mirror| 100daysofhope| bodies| image| days
10	158	0.89	Ice-Cream	ice| cream| chocolate| snack| chocolateporn| icecream| cookies
11	157	0.88	Vegan Fitness	cancersurvivor| cleanandsober| veganfit| fitover45| veganstrong| fitnessathome| veganaf

Note. Topic -1 indicates the *Outliers*, tweets that did not relate to any of the final topics.

Table C2

Top 11 most relevant Pro-anorexia topics

Topic	Count	% of Total	Label	Topic Words
-1	120003	41.56	Outliers	edtwt| thinspo| thinspiration| skinny| im| ana| proana
0	24999	8.66	Thinspiration	shes| perfect| thinspiration| body| thinspo| love| perfection
1	11220	3.89	PA Messages	eat| proana| food| starve| promia| dont| eating
2	7762	2.69	Leg Thinspiration	legs| thighs| thighgap| perfect| fat| thinspo| want
3	6064	2.10	Anasister	anasisters| anasister| follow| love| anafamily| thanks| lt3
4	5720	1.98	Bones	bones| chest| hunger| collarbones| people| proana| ached
5	5468	1.89	Thinspiration Trends	thinspiration| need| tbt| fitspiration| look| throwback| like
6	5420	1.88	Gym Inspiration	gym| thinspiration| workout| run| fitness| healthy| lunch
7	5309	1.84	Weight Loss	lost| weight| pounds| lose| lbs| scale| bmi
8	4838	1.68	Fasting	fast| fasting| hours| hour| day| days| starting
9	4659	1.61	Mutuals	moots| new| mutuals| cw| sheher| gw| ugw
10	3986	1.38	Skinny Outfits	wear| jeans| dress| shorts| look| cute| clothes
11	3486	1.21	Weight Progress	httpstcou3ufes6wtj| weightloss| months| years| progress| year| face

Note. Topic -1 indicates the *Outliers*, tweets that did not relate to any of the final topics.

Appendix D

ChatGPT Prompt

```
I have a topic that contains the following documents: [DOCUMENTS]
The topic is also described by the following keywords: [KEYWORDS]
Based on the information above, extract a short topic label in the
following format: topic: <topic label>
The topic label should be as short as possible while being as
descriptive as possible but no more than 4 words in length and no
two topics should have the same label.
```