Article

The Emotional Landscape of Technological Innovation: A Data-Driven Case Study of ChatGPT's Launch

Lowri Williams * and Pete Burnap

School of Computer Science & Informatics, Cardiff University, Cardiff CF24 4AG, UK; burnapp@cardiff.ac.uk
* Correspondence: williamsl10@cardiff.ac.uk

Abstract

The rapid development and deployment of artificial intelligence (AI) technologies have sparked intense public interest and debate. While these innovations promise to revolutionise various aspects of human life, it is crucial to understand the complex emotional responses they elicit from potential adopters and users. Such findings can offer crucial guidance for stakeholders involved in the development, implementation, and governance of AI technologies like OpenAI's ChatGPT, a large language model (LLM) that garnered significant attention upon its release, enabling more informed decision-making regarding potential challenges and opportunities. While previous studies have employed data-driven approaches towards investigating public reactions to emerging technologies, they often relied on sentiment polarity analysis, which categorises responses as positive or negative. However, this binary approach fails to capture the nuanced emotional landscape surrounding technological adoption. This paper overcomes this limitation by presenting a comprehensive analysis for investigating the emotional landscape surrounding technology adoption by using the launch of ChatGPT as a case study. In particular, a large corpus of social media texts containing references to ChatGPT was compiled. Text mining techniques were applied to extract emotions, capturing a more nuanced and multifaceted representation of public reactions. This approach allows the identification of specific emotions such as excitement, fear, surprise, and frustration, providing deeper insights into user acceptance, integration, and potential adoption of the technology. By analysing this emotional landscape, we aim to provide a more comprehensive understanding of the factors influencing ChatGPT's reception and potential long-term impact. Furthermore, we employ topic modelling to identify and extract the common themes discussed across the dataset. This additional layer of analysis allows us to understand the specific aspects of ChatGPT driving different emotional responses. By linking emotions to particular topics, we gain a more contextual understanding of public reaction, which can inform decision-making processes in the development, deployment, and regulation of AI technologies.

Keywords: emerging technologies; technology barriers; technology adoption; emotive language analysis; topic modelling; chatgpt

1. Introduction

In an age where technology and its advancements redefine the boundaries of human–machine interactions, the launch of new technological tools, especially those with the potential to revolutionise communication, often elicits a spectrum of public reactions [1]. Such reactions can offer crucial guidance for stakeholders involved in the development,

implementation, and governance of such technologies, providing valuable insights into user acceptance, potential barriers, and opportunities for improvement [2].

Recent studies (e.g., [3–5]) have explored the application of data-driven and text mining approaches to extract information about public reactions from social media platforms (e.g., Twitter, now formally known as 'X'), leveraging the wealth of public discourse surrounding emerging technologies. However, these works have predominantly focused on extracting sentiment polarity, which categorises responses as positive or negative. Although this binary approach provides a basic indication of adoption or acceptance, it fails to capture the nuanced emotional landscape surrounding technological adoption.

To overcome this limitation, this paper presents a comprehensive analysis to investigate the emotional landscape surrounding technology adoption. This approach goes beyond sentiment polarity, allowing specific emotions to be pinpointed, such as excitement, surprise, frustration, and others that users experience with the emergence of new technologies. This granular emotional analysis offers a more sophisticated understanding of public reactions, enabling more informed decision-making in the development and deployment of technology.

In recent years, artificial-intelligence-(AI)-powered technologies have been widely adopted in various fields, including health care, finance, transportation, and entertainment [6], as well as being integral to our daily lives (e.g., virtual assistants, personalised recommendations, facial recognition). One of the most recent breakthroughs in AI research has been the emergence of large language models (LLMs) that are built on generative pre-trained transformers (GPTs) [6]. In particular, OpenAI's ChatGPT [7] emerged as a groundbreaking innovation in the domain of conversational AI in November 2022. This AI-powered chatbot quickly gained widespread attention for its ability to engage in human-like conversations and its unprecedented capabilities in natural language processing, including answering questions, providing explanations, assisting with writing tasks, and even generating code. Given the advantages of ChatGPT over traditional chatbots, it attracted more than 1 million users in its first week after it was launched, surpassing adoption rates of other popular online platforms such as Netflix, Facebook, and Instagram [4].

However, along with its rapid popularity, ChatGPT has prompted concerns about the broader implications of its use [6]. Among these concerns is the potential for its exploitation for malicious purposes [8], the possibility of the technology exacerbating preexisting societal biases by inadvertently reflecting biases from its training data, causing ChatGPT to produce biased responses [9], and its ability to produce highly convincing fake text.

Given these two contrasting scenarios—on the one hand, the rapid adoption and widespread benefits of ChatGPT, and on the other, the concerns around its potential misuse and unintended societal impacts—it becomes crucial to explore how users emotionally respond to this technology. These varied reactions highlight the complexity of public sentiment toward AI systems. As such, in this paper, we focus on ChatGPT as a case study, aiming to capture and analyse the full spectrum of emotions expressed by users during its release to the public. By doing so, we ensure that we explore the emotional landscape in all its nuances, offering a deeper understanding of both the enthusiasm and apprehension surrounding its adoption and integration.

The main contributions of the work presented herein are as follows:

- A deep exploration of the emotional responses of early technology adopters. This study goes beyond traditional sentiment analysis by categorising user responses into specific emotions. This approach pushes the boundaries of sentiment analysis, offering a more nuanced and fine-grained understanding than the conventional positive, negative, and neutral classifications.

- A comprehensive analysis of key topics driving emotional responses among early technology adopters. By mapping user discussions to distinct emotional categories, this analysis uncovers the underlying reasons behind the emotions expressed. This offers valuable insights into the factors influencing both positive and negative reactions, providing a deeper understanding of user sentiment towards new technologies.

This study was designed as shown in Figure 1: (1) compile a corpus of texts containing references to ChatGPT during its launch, (2) preprocess text responses using traditional natural language processing (NLP) techniques, (3) divide texts based on their publication date, (4) for each dataset in (3), automatically extract emotions expressed within text segments, (5) for each emotive category in (4), apply a topic modelling approach to automatically identify aspects discussed in text segments, and (6) visualise and analyse the results.
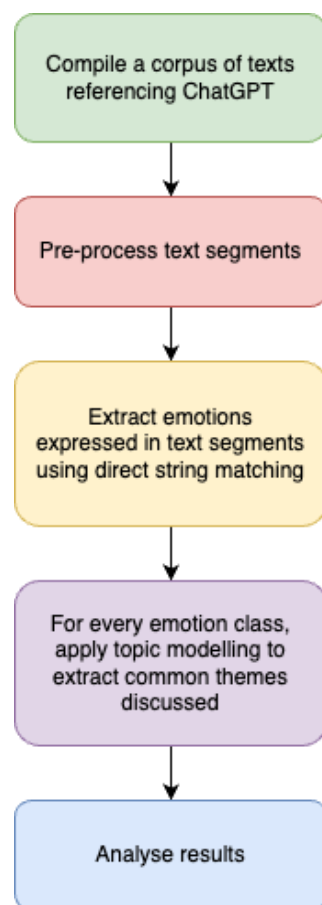
Figure 1. An overview of the study design.

The remainder of this paper is structured as follows: Section 2 presents the related work. Section 3 discusses the collection of texts used to support the experiments herein and the techniques used to prepare the data for such experiments. Sections 4 and 5 discuss aspect-based emotive analysis and how it was applied to the datasets. Section 6 presents and discusses the results. Section 7 concludes the paper; finally, Section 8 discusses future work.

2. Related Work

Studies investigating technological adoption have traditionally employed various methodological approaches, ranging from theoretical frameworks to qualitative and quantitative analyses. Theoretical models like the Technology Acceptance Model (TAM) and Unified Theory of Acceptance and Use of Technology (UTAUT) have been instrumental in

understanding factors influencing technology uptake by examining perceived usefulness and ease of use [10,11].

While early research predominantly used survey-based and interview methodologies, recent studies have turned to text mining approaches to automatically analyse and help understand the factors influencing consumer adoption or rejection of emerging technologies from a large landscape of adopters [3]. These approaches offer advantages in processing large volumes of user-generated content, particularly from social media platforms. Most recently, Williams et al. [3] introduce a scalable and automated framework designed to monitor the potential adoption or rejection of emerging technologies by analysing social media discourse. They compiled a substantial corpus of social media texts referencing new technologies and applied text mining techniques to extract sentiments related to various technological aspects. Powered by sentiment analysis, they hypothesised that positive sentiments correlate with a higher likelihood of acceptance of technology, while negative sentiments indicate potential rejection. To validate this hypothesis, a ground truth analysis was performed comparing the automated sentiment extraction results with human annotations. The findings demonstrated that the automated approach effectively captures sentiments comparable to human evaluations, providing valuable information on user behaviours, emerging demands, and uncertainties surrounding new technologies.

Others have applied sentiment analysis across diverse technological domains. Kwarteng et al. [5] investigated consumer perceptions of autonomous vehicles through Twitter data, while Efuwape et al. [12] examined digital collaborative tool adoption using sentiment analysis. Hizam et al. [13] also explored technology adoption behaviours by analysing social media content, focusing on factors like perceived utility and social impact. Mardjo and Choksuchat [14] and Caviggioli et al. [15] used sentiment analysis to examine public perceptions of Bitcoin, aiming to forecast market behaviour and understand adopter reactions.

While existing research employs sentiment analysis to understand technology adoption patterns, such approaches that focus solely on sentiment polarity (positive/negative) provide limited insights into the nuanced factors driving adoption reactions and decisions. Basic sentiment analysis fails to capture the true nuances behind the specific aspects of technology that users embrace or reject, leaving a critical gap in our understanding of adoption behaviours. In this case, studies such as Lee and Kang [16] apply topic modelling to analyse technology and innovation management research, identifying key themes and trends over two decades. Their study highlights a shift from traditional management theories to data-driven, empirical research, with growing interest in technology adoption, patent analysis, and social network analysis. Similarly, Cobelli and Blasi [17] produce a similar topic modelling and bibliometric analysis to understand the evolution of technological innovation adoption in the healthcare industry. However, these works focus on academic discourse, lacking insight into public perception, emotional responses, and real-world adoption barriers.

This study presents a comprehensive framework for emotional landscape analysis, addressing the limitations of previous research by integrating emotional analysis with topic modelling. By leveraging text mining techniques on large-scale social media corpora, the research moves beyond polarised sentiment analysis to extract nuanced emotional reactions, such as excitement, fear, surprise, and frustration, providing deeper insights into technological acceptance and long-term societal impacts. This approach has significant implications for technology developers and policymakers, enabling them to optimise product features, address adoption concerns, and align innovation strategies with user expectations. Unlike other research that primarily serves academia and policy, this case study offers an adaptable methodology for analysing technology adoption across various

domains. By combining emotion extraction with topic modelling, the research bridges the gap between technology development and public perception, supporting more informed decision-making in innovation strategy, deployment, and governance.

3. Text Corpus

To assess the public sentiment of early ChatGPT adopters, in this paper, social media data was collated, specifically from Twitter, a social networking service that enables users to send and read tweets, text messages consisting of up to 280 characters. Recent shifts in Twitter's data collection policies, under the platform's rebranded identity as 'X', have significantly impacted the social media platform's data accessibility for research purposes. In response to these data accessibility limitations, other rich and well-established resources in the data science community, such as Kaggle (https://www.kaggle.com, accessed on 5 May 2024), host an array of relevant datasets, including compilations of Twitter conversations.

To gather relevant data for this study, a comprehensive search was conducted on Kaggle using specific keywords related to the research focus. The search criteria included datasets of tweets that had been collected using references to 'ChatGPT,' 'GPT3,' 'GPT4,' and 'OpenAI.' This targeted approach allowed us to identify and compile a rich corpus of Twitter conversations specifically centred around ChatGPT and related AI technologies. By leveraging these pre-collected datasets, we were able to circumvent the challenges posed by Twitter's new data policies while still accessing a robust and relevant sample of public discourse on the topic.

Table 1 reports the distribution of tweets across each dataset. Most datasets were collected using Twitter's API, typically by targeting specific keywords or hashtags related to ChatGPT. However, in many cases, the detailed methodology behind data collection is not fully documented, which can hinder reproducibility and transparency. Additionally, none of the datasets explicitly filter tweets by region or age group. Although this broad approach enables the capture of a wide range of opinions, it can also introduce bias, particularly the over-representation of English-speaking users or individuals from regions with greater Twitter usage.

All datasets are publicly available on Kaggle, making them accessible for further research and comparative analysis. To mitigate potential biases and enhance the robustness of the study, researchers are encouraged to consider data augmentation by incorporating datasets from more diverse sources or languages. If metadata such as user location or language preferences is available, it should be analysed to evaluate and correct for regional imbalances.

Table 1. Description of the text corpus collated from Kaggle.

Dataset	Number of Tweets	Keywords	Date Range
[18]	500,002	ChatGPT, Chat GPT	04/01/2023–29/03/2023
[19]	1000	Unknown	28/02/2023
[20]	43,004	ChatGPT, GPT3, GPT4	03/04/2023–12/05/2023
[21]	98,759	ChatGPT	18/03/2023–21/03/2023
[22]	28,710	GPT4	14/03/2023–12/04/2023
[23]	50,001	ChatGPT	22/01/2023–24/01/2023
[24]	478,301	ChatGPT	30/11/2022–08/04/2023
[25]	219,294	ChatGPT	30/11/2022–31/12/2022
[26]	21,514	Unknown	05/01/2023–19/01/2023
[27]	3,821,843	Unknown	30/11/2022–14/05/2023

Following the concatenation of each dataset, using Python's regular expression package, RegEx (version 2020.9.27), retweets were removed by identifying text that began with the standard Twitter retweet syntax (RT @username:), which typically signifies a reposted tweet. By filtering out such entries using a regular expression pattern (e.g., RT @), the dataset retains only original content posted each day. This helps reduce redundancy and the disproportionate influence of highly retweeted content. Duplicate tweets collected each day according to their posted timestamp were then removed. The remaining dataset contained a total of 5,262,428 tweets. The dataset spans between 30 November 2022 and 14 May 2023 (167 days), with 20 March 2023 reporting the most tweets (9171) and 23 April 2023 reporting the fewest (54).

The data preparation and analysis in this study were conducted using Python (version 3.7.2). For text preprocessing, the following traditional NLP techniques were applied:

- Converting text to lowercase.
- Removing additional white spacing, mentioned usernames, hashtags, URLs, punctuation, digits, and emojis using Python's regular expression package, RegEx (version 2020.9.27).
- Tokenising text using Python's natural language package, Natural Language Toolkit (NLTK) (version 3.4.1).
- Removing stop words (e.g., this, are, and a) as part of the NLTK package.
- Stemming tokens using the Porter Stemmer as part of the NLTK package.
- The keywords (i.e., 'ChatGPT', 'GPT3', 'GPT4', and 'OpenAI') used to search for tweets were removed, as their presence may disproportionately influence future topic modelling results. Such terms may dominate the text due to their frequency and centrality, not necessarily because they reflect the true variety of subtopics within the data. By removing them, topic distributions have a stronger likelihood of emerging based on semantic context rather than being skewed by the repetitive presence of such words. Furthermore, this practice supports a more unbiased exploration of the content landscape, uncovering hidden or less obvious themes such as sentiment, application domains, concerns, or trends related to ChatGPT, rather than reaffirming its mention.

4. Extracting Emotions from Text

Extracting emotions from text presents a significant challenge in natural language processing, primarily due to the lack of consensus on which terms should be classified as being emotive [28]. This ambiguity has led to the development of various emotional classification schemes, each structuring emotion categories differently. Such schemes range from Ekman's [29] six basic emotions to more complex models like Plutchik's [30] wheel of emotions, illustrating the diverse approaches to emotion categorisation in text analysis.

Among the advancements in emotive text analysis, emotive hierarchies (e.g., [31–33]) have emerged as a powerful tool for capturing the complexity of human emotions expressed in text. Emotive hierarchies are structured frameworks that organise emotions into multiple levels, typically starting with broad sentiment categories (such as positive or negative) and progressively narrowing down to more specific emotional states such as happiness, sadness, love, and anger [28]. At the most granular level, such emotions are further refined into nuanced states that align with their overarching polarities, such as optimism (a form of happiness), misery (a form of sadness), and passion (a form of love).

Such hierarchies offer several advantages over traditional binary or categorical emotion classification methods. Firstly, they provide a more granular representation of emotions, allowing for the detection of subtle nuances in emotional expression. Secondly, the hierarchical structure reflects the natural relationship between emotions, capturing how certain emotions are related or can co-occur. This approach enables researchers and practitioners

to analyse text at various levels of emotional specificity, from general sentiment to precise emotional states, thereby offering a more comprehensive understanding of the affective content in textual data.

Inspired by Williams et al. [34], who investigate and compare the utility of different emotional classification schemes from both a machine and human perspective, in this paper, WordNet-Affect [35] was employed. Building upon the foundation of WordNet, a comprehensive lexical database for the English language that organises nouns, verbs, adjectives, and adverbs into interconnected sets of synonyms called synsets [36], WordNet-Affect was developed as a specialised lexical model to classify both direct (e.g., joy, sad, happy) and indirect (e.g., pleasure, hurt, sorry) expressions of affect. Since its creation, it has become a valuable lexical resource, supporting numerous natural language processing studies, particularly in the field of sentiment analysis (e.g., [34,37,38]).

Figure 2 shows an excerpt from the WordNet-Affect hierarchy. The local version of the lexicon used in this paper consists of 1532 emotion terms, including all derivational and inflectional forms of the 798 word senses originally found in WordNet-Affect. Table 2 shows the distribution of emotions across the seven levels of WordNet-Affect's hierarchical structure, as well as examples of emotions at each level.
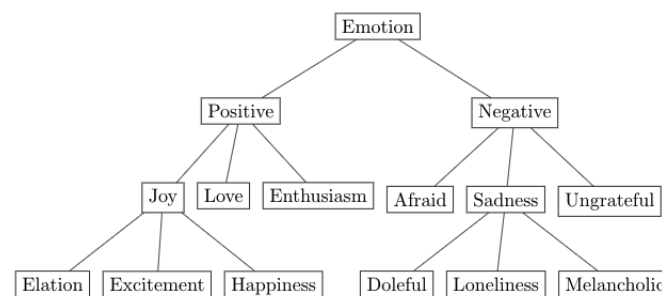
Figure 2. An excerpt from the WordNet-Affect hierarchy [34].

**Table 2.** Distribution of emotions across the WordNet-Affect hierarchy.

Level	Examples of Emotions	No. of Emotions
1	Positive, Negative, Neutral, Ambiguous	4
2	Love, Surprise, Sad, Fear	162
3	Adore, Amusement, Distress, Embarrassed	582
4	Jolly, Peaceful, Annoyed, Confused	527
5	Pride, Fulfilment, Aggressive, Livid	198
6	Remorse, Contrite, Guilt, Woeful	46
7	Covet, Jealousy, Greedy, Green-eyed	12

To automatically extract a breadth of emotions, a direct string matching approach was applied to automatically extract emotive terms from WordNet-Affect present in tweets. Table 3 reports examples of tweets mapped to emotions. More specifically, a total of 4,206,691 tweets (79.9% of the original dataset) did not contain emotion and, therefore, were omitted from the dataset. A total of 1,055,737 tweets (20.1% of the original dataset) contained an expression that could be mapped to emotions from the WordNet-Affect lexicon, with 207,928 of those tweets containing one emotion. For the 847,809 tweets containing more than one emotion, they were represented in multiple emotion categories. This ensures that the complex and multifaceted emotional expressions found in the tweets are accurately

captured and reflected in the analysis, allowing for a more nuanced understanding of the emotions conveyed in the dataset.

However, it is important to acknowledge that relying solely on a lexicon-based, string-matching technique may introduce limitations. The informal, creative, and evolving nature of social media language means that many emotional expressions, particularly those conveyed through slang, sarcasm, or implied sentiment, may not be captured, potentially resulting in false negatives or misclassification. A lexicon-based approach may be validated or complemented with machine learning classifiers fine-tuned for emotion detection, which may be equipped to handle contextual nuances and linguistic variation.

To analyse emotions more effectively, fewer emotive categories are needed. In this case, WordNet-Affect's hierarchical structure can be leveraged, where emotive terms can be traced back to their ancestral emotion within the second level of the hierarchy. For example, in Figure 2, "Happiness" and "Excitement" can be categorised as "Joy", and "Doleful" and "Loneliness" can be categorised as "Sadness". This categorisation approach resulted in 33 distinct emotion groups. Figure 3 reports the emotion categories as well as the distribution of tweets across them.
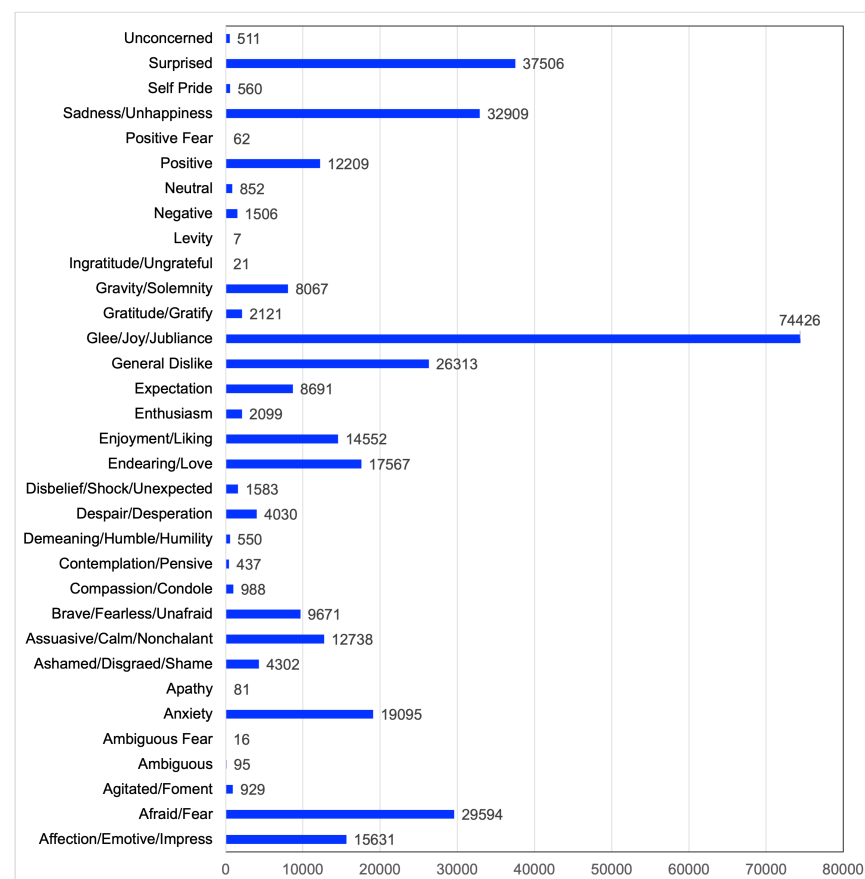
Figure 3. Distribution of tweets across 33 emotion categories.

To better illustrate the emotion detection process and its contextual grounding, Table 3 presents a series of tweets annotated with both their original emotion (as derived from the WordNet-Affect lexicon) and a more generalised emotional category. These examples serve to demonstrate how emotions were identified and interpreted within the discourse around ChatGPT. The first tweet reflects anxiety surrounding AI's impact on academia, specifically regarding the rapid deployment of AI-detection tools and their perceived limitations. The emotion "worried" was mapped to the broader category "Anxiety" to capture this collective unease about educational integrity. The second tweet highlights a user comparison between

two AI outputs, where the author notes a "sad practical answer" from ChatGPT. This sentiment is directly identified as "Sad" and categorised under "Sadness/Unhappiness", showing how user expectations of AI outputs can evoke emotional disappointment. In the third example, the user expresses affection toward ChatGPT while humorously questioning their own intelligence. The phrase "I love ChatGPT" is associated with the original emotion "Love", which was generalised as "Endearing/Love" to encapsulate both appreciation and a self-deprecating tone. Finally, the fourth tweet conveys enthusiasm about ChatGPT's potential in retail, portraying optimism and forward-looking excitement. The original term "Excited" was generalised to "Glee/Joy/Jubliance", reflecting broader positive anticipation in commercial contexts.

Table 3. Examples of tweets mapped to emotions.

Tweet	Original Emotion	Generalised Emotion
The plagiarism detector will introduce its AI detection tool tomorrow, hoping to protect academic integrity in a post-ChatGPT world. The speedy launch and lack of an opt-out have academics worried.	Worried	Anxiety
First one is ChatGPT from OpenAI. Second one is ours. The first one gives a sad practical answer.	Sad	Sadness/Unhappiness
I love ChatGPT but I'm dumber now if that's even possible.	Love	Endearing/Love
Just read an amazing article about how ChatGPT is revolutionizing the retail industry and changing the way we shop. We're excited to see the endless possibilities this technology can bring to the shopping experience!	Excited	Glee/Joy/Jubliance

5. Extracting Topics from Emotive Texts

Following the extraction of emotions from tweets, to add an additional layer of analysis, topic modelling is integrated into the framework. This allows us to contextualise the emotional responses about the specific topics being discussed, providing a richer, more comprehensive view of public sentiment and the factors influencing the adoption of AI technologies like ChatGPT.

Topic modelling serves as a crucial technique to uncover latent semantic structures—commonly referred to as "topics"—from large text corpora [39,40]. It has been successfully applied in domains ranging from online social networks [41,42] to healthcare [43] and mental health [44].

There are various methods by which texts can be distributed into topics. Such methods vary in complexity, the representation of text segments taken as the model's input, their computational speed, and their performance. However, such approaches often yield topics as collections of isolated words, which can lack clarity and semantic coherence. While these word lists suggest key themes, they may not fully capture the context or relationship between terms, making it challenging to interpret the exact meaning of the topics or their relevance to user discussions.

In this case, to ensure high interpretability and consistency, we adopted the enhanced topic modelling pipeline proposed by Williams et al. [40], which is accompanied by publicly

available and reproducible code. This pipeline builds on the Latent Dirichlet Allocation (LDA) algorithm and includes preprocessing, topic generation, and a post-processing stage designed to convert token lists into more semantically coherent topic descriptions. For example, instead of keywords like 'AI,' 'job,' and 'disruption,' the method produces descriptors such as "concerns about AI-driven job disruption," offering clearer insights into public narratives. For more detail as to how such topic representations are generated, see [40].

The preprocessing steps include:

- Lowercasing text
- Removing punctuation, digits, white space, and emojis
- Tokenisation
- Stop-word removal
- Stemming

Topic modelling is performed using the 'LatentDirichletAllocation' class from Scikit-learn [45], following the same configuration as the original implementation in [40]. The LDA model was applied in a plug-and-play fashion using the default hyperparameters:

- number of topics = 10,
- random state = 10
- compute perplexity ever n iterations = -1
- No manual specification of alpha or beta (i.e., doc_topic_prior and topic_word_prior left as default)

The choice to fix the number of topics at ten across all emotional categories per day was based on maintaining consistency with the original methodology, rather than conducting dataset-specific optimisation (e.g., via coherence or perplexity-based tuning). While perplexity is a commonly used metric for assessing topic model performance, it has been shown to not always align with human interpretability [46]. Given the focus on cross-category comparison and practical interpretability, a stable and comparable topic structure was prioritised over per-category optimisation. It is also worth noting that multiple runs were not conducted to assess topic stability, as the goal of this step was not model tuning but to support the downstream emotion–topic alignment in line with the original reproducible framework.

Overall, this enhanced topic modelling approach offers an efficient and consistent method for extracting interpretable themes from large-scale social media text, supporting a more nuanced understanding of emotional discourse around ChatGPT and similar technologies.

To ensure consistency and comparability across emotion categories, a fixed topic count strategy was adopted. Specifically, for each of the 33 emotional categories described in Section 4, topic modelling was applied to extract ten topics per day from the corresponding tweets.

6. Results and Discussion

The chronological analysis depicted in Figure 4 underscores the shift in emotional sentiment towards ChatGPT over time. Each row represents a unique generalised emotional category (e.g., Affection/Emotive/Impress, Endearing/Love, Sadness/Unhappiness), while each column represents a single day from December 2022 to May 2023. The values are normalised per day, such that the total number of emotion-tagged tweets on any given day sums to 1. This normalisation enables direct comparison across days, regardless of tweet volume. Although colours are used solely to differentiate emotion categories (i.e., no specific colour-to-emotion mapping exists), darker shades within each row indicate a

higher relative proportion of that emotion on that day. For example, a darker cell in the "Anxiety" row on a given date signifies that anxiety-related tweets made up a larger share of total emotional tweets that day. This visualisation facilitates the identification of key emotional spikes or transitions in public discourse surrounding ChatGPT.

The heatmap allows emotional sentiments across different periods to be contrasted and compared, offering insights into the public's evolving perceptions and attitudes. For example, the "Surprised" and "Enjoyment/Liking" categories initially follow a consistent pattern where a high volume of tweets about ChatGPT were shared but showed a marked increase in volume a few days post-launch, coinciding with the period when users began interacting with ChatGPT. This increase may capture users' reactions to experiencing ChatGPT's capabilities firsthand, which perhaps exceeded expectations or revealed the unforeseen abilities of the AI, prompting a heightened collective response of surprise and satisfaction in the discourse.

Similarly, the "Anxiety" category displayed a consistently significant emotional reaction throughout the timeline. However, this sustained intensity likely reflects the public's apprehensions surrounding the launch of ChatGPT. Given the transformative potential of such AI technology, it is plausible that discussions were tinged with concern over its implications for privacy, employment, and the broader societal impact, leading to heightened expressions of anxiety across the observed period. This aligns with Moore's [47] adopter categories, where early adopters engage with the innovation enthusiastically while the early and late majority remain hesitant due to perceived risks.

Other emotional categories, such as "Afraid/Fear" and "Agitated/Foment," displayed a generally lower volume of tweets over the timeline. However, a noticeable increase in tweet volume occurred a few days after the release of ChatGPT. This increase could potentially be attributed to users' initial apprehension and heightened emotional responses as they began to grapple with the AI's capabilities and implications, confronting the tangible reality of this advanced technology's role in their digital lives. Such sentiments reflect the resistance typically found among laggards in the technology adoption lifecycle, who are more risk-averse and cautious about adopting new innovations [48].

Figure 5 presents a matrix illustrating a subset of results following the extraction of topics (discussed in Section 5), showing how specific emotional categories are distributed across identified topics. Each row corresponds to a distinct topic label, and each column represents one of the 33 emotional categories. A coloured cell indicates that a given emotion was detected in tweets associated with that topic. Colour hues are used only to visually distinguish between categories and do not represent intensity or frequency.

The analysis reveals a complex landscape of discourse. A diverse range of topics emerged, spanning from technological capabilities to ethical considerations and potential applications across multiple sectors, underscoring ChatGPT's far-reaching implications. Positive emotional categories, such as "Endearing/Love" and "Affection/Emotive/Impress," tied to topics like "AI Generated content" and "AI Writing Tool Produce Good Blog Content," indicate strong user endorsement and satisfaction with ChatGPT's creative and practical applications. These positive sentiments are likely to drive further adoption and integration of ChatGPT in various domains, as users see tangible benefits and improvements in their tasks. This pattern aligns with the early adopters and early majority in the adoption lifecycle, who embrace innovations that demonstrate clear advantages.
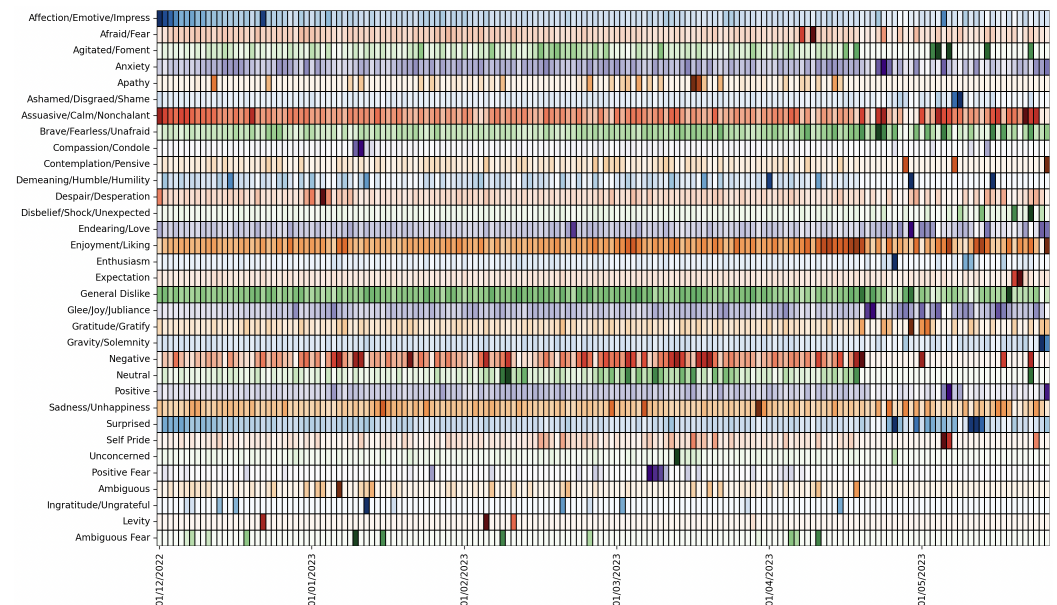
Figure 4. Chronological heatmap of emotional expressions in ChatGPT-related tweets (December 2022–May 2023). Each row represents a distinct emotional category, and each column corresponds to a single day. The data has been normalised so that the number of tweets in each emotion category per day is divided by the total number of emotional tweets on that day, allowing for day-to-day comparison regardless of tweet volume. Colours are used to distinguish emotions, while darker colour intensity indicates a higher relative frequency of that emotion on a given day.

The highly polarised sentiment where the potential for AI to replace human roles is met with both apprehension and concern reflecting societal fears of job displacement and the ethical implications of AI advancement. It is also met with acceptance and excitement for the efficiency and innovation AI can bring, underscoring the dual nature of public perception. For example, the topics "AI Tool Like ChatGPT Might Disrupt", "AI Writes Better Online Discussion Response", and "Change Health Care Forever" were extracted from texts reflecting the emotional categories "Afraid/Fear", "General Dislike", and "Despair/Desperation".

Whereas the topics "AI Replace Humankind", "AI Tool Developing Consumer Trust", and "Revolutionizing Industry Like Healthcare" were extracted from texts predominantly evoking the positive emotional categories "Brave/Fearless/Unafraid", "Positive", and "Enjoyment/Liking", which denote the recognition of the practical benefits and enhancements that ChatGPT can bring to various professional fields, highlighting an optimistic outlook on AI's role in enhancing productivity and skill development.

This duality highlights the need for careful management and communication from AI developers and policymakers to balance these concerns with the promising benefits, ensuring a transition that maximises positive outcomes while addressing and mitigating fears. The evolution of sentiment, as shown in the data, suggests that ChatGPT's adoption trajectory follows Rogers' [48] stages of technology adoption, where initial scepticism gradually gives way to broader acceptance as users become more familiar with the technology's capabilities and benefits.

The diverse emotional responses to ChatGPT-related topics reflect the complex landscape of AI adoption. While there is significant enthusiasm and positive anticipation about the capabilities and benefits of ChatGPT, there is also notable concern regarding its reliability, ethical implications, and security. This duality in sentiment underscores the need for ongoing dialogue, transparent practices, and continuous improvements in AI technologies. Ensuring that AI developments align with user expectations and address

their concerns will be key to fostering a balanced and widespread adoption of AI tools like ChatGPT, ultimately progressing through the technology adoption lifecycle towards mainstream acceptance.

Figure 5. This figure presents an excerpt of the topics identified through topic modelling applied to tweets containing emotional expressions. Each row corresponds to a unique topic, while each column represents a generalised emotional category. A coloured cell indicates that the corresponding topic appeared within that emotion category; colour hue is used only to differentiate emotions visually and has no analytical significance.

The emotion concurrence matrix, shown in Figure 6, captures how often different emotional categories co-occur across various topics discussed in ChatGPT-related tweets. Each cell represents the number of topics in which a pair of emotions appear together, allowing us to explore patterns of emotional entanglement within thematic content. The diagonal values, representing self-co-occurrence (e.g., "Joy" with "Joy"), have been omitted for clarity, as they reflect how frequently each individual emotion appears across topics.

The emotion concurrence matrix reveals several notable patterns in how emotional categories co-occur in ChatGPT-related tweets. Emotions rarely appear in isolation; instead, they tend to cluster in meaningful ways, indicating the complex nature of public sentiment toward generative AI. Among the most frequent co-occurrences are those involving "General Dislike", which appears prominently alongside "Anxiety" and "Ashamed/Disgraced/Shame," which may reflect underlying societal discomfort or ethical unease surrounding AI use, particularly in sensitive domains like education or employment. Emotions such as "Assuasive/Calm/Nonchalant" and "Brave/Fearless/Unafraid" also show multiple connections, often appearing in contrastive contexts—tweets that balance optimism with caution or confidence with critique. Similarly, "Enjoyment/Liking" often co-occurs with "Endearing/Love" and "Surprised", reflecting consistent patterns in positive affect when users express satisfaction or excitement about ChatGPT.

In contrast, several emotion categories demonstrate minimal co-occurrence with others. These include "Ambiguous Fear", "Ingratitude/Ungrateful", "Positive Fear", and "Levity", which may indicate their rarity in the dataset or challenges in detecting such nuanced sentiments using a lexicon-based approach. The relative isolation of these categories might also suggest that they tend to occur in more context-specific or subtle expressions that are not easily captured through surface-level keyword matching.

Overall, the results highlight clusters of emotion that align with broader sentiment dynamics—positive categories reinforcing each other in celebratory or promotional content and negative or anxiety-related emotions clustering around moments of uncertainty, fear, or disillusionment. These findings not only underscore the multidimensional emotional landscape surrounding ChatGPT but also point to the value of co-occurrence analysis in uncovering how emotional expressions are entangled in public discourse.
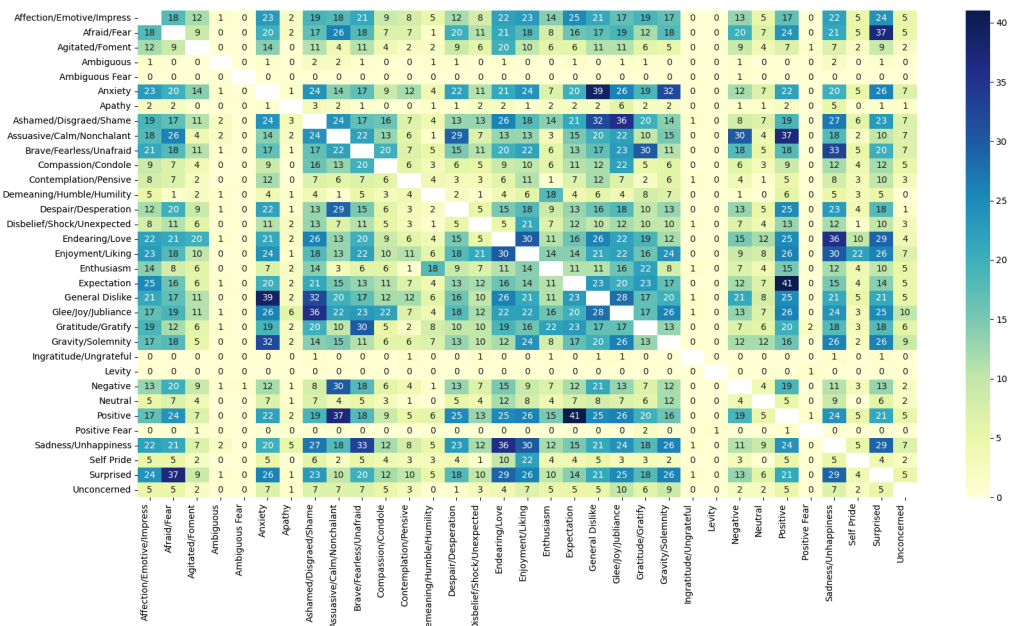
Figure 6. Emotion concurrence matrix showing co-occurrence of emotional categories across ChatGPT-related topics.

7. Conclusions

This paper presents a comprehensive analysis for investigating the emotional landscape surrounding the adoption of emerging technologies, using the launch of OpenAI's ChatGPT as a case study. By utilising text mining techniques on a large corpus of social media posts, emotions such as excitement, fear, frustration, and surprise were extracted and

categorised. Unlike traditional sentiment analysis, which often reduces public reactions to a binary polarity, this approach captures the nuanced emotional responses that better reflect the complex dynamics involved in technology adoption. In addition to emotional analysis, we introduced a temporal dimension that tracks how such emotions evolved over time, providing insights into the shifting public sentiment as users become more familiar with the technology.

By applying topic modelling to uncover features or events that trigger particular emotional responses, further light is shed on how emotions are tied to key aspects of ChatGPT, such as its technological capabilities and ethical implications. Topics related to AI-generated content and the practical applications of ChatGPT in areas like writing and healthcare elicited positive emotional categories such as "Endearing/Love" and "Affection/Emotive/Impress," signalling strong user endorsement and satisfaction. These positive emotional responses are likely to drive further adoption of ChatGPT, as users perceive tangible benefits in improving their daily tasks and professional activities.

Conversely, more polarised topics, such as the potential for AI to replace human roles, prompted mixed emotions, including fear and excitement. While some users expressed apprehension about job displacement and the ethical ramifications of AI advancement, others embraced the efficiency and innovation AI brings. This duality in sentiment highlights both the promise and the challenges of AI adoption, underscoring the need for thoughtful communication and policy interventions to address societal fears while promoting the potential benefits of AI.

The findings emphasise the importance of carefully managing the public discourse around AI technologies, particularly in addressing concerns related to security, ethics, and reliability. As the emotional responses suggest, successful AI adoption hinges not only on demonstrating the practical advantages of these technologies but also on ensuring that developers and policymakers engage in transparent practices that align with user expectations and mitigate concerns. This dual focus on maximising positive outcomes while addressing the public's fears will be key to fostering a balanced, widespread adoption of AI tools like ChatGPT.

8. Future Work

Emotions, such as fear, excitement, frustration, and curiosity, are critical factors that shape users' trust in AI technologies like ChatGPT. Trust is often a key determinant of whether individuals are willing to adopt and integrate new technologies into their daily lives. For example, fear of the unknown can create resistance to AI, particularly when users feel that they lack control or understanding of how the technology operates. On the other hand, excitement can drive curiosity and early adoption, but it may fade if users encounter frustrations such as inaccuracies or ethical concerns over data usage. Based on the work presented herein, future research directions include exploring how specific emotional responses influence the development and maintenance of trust over time. A deeper understanding of how positive emotions (like excitement or relief when AI solves a problem) and negative emotions (like frustration with AI errors) impact trust can guide the design of AI systems that foster sustained user confidence. Additionally, research could investigate strategies for mitigating negative emotions, such as improving transparency or offering clearer explanations for AI decision-making, thereby strengthening user trust and promoting wider acceptance of AI technologies in sensitive sectors such as healthcare, finance, and education. This line of inquiry could have significant implications for both user experience design and AI policy development.

An important direction for future research involves incorporating external event timelines to better contextualise emotional fluctuations. While this study focused on data-

driven patterns of emotional expression over time, future analyses could align emotional spikes with key milestones such as OpenAI announcements, product launches (e.g., GPT-4), or major media coverage. This would allow for more targeted exploration of causal or correlational relationships between real-world events and public sentiment, thereby strengthening the interpretive power of temporal emotion analysis. Combining this with entity recognition or event tagging could further refine the connection between specific developments and emotional responses.

References

1. Bauer, M.W. Public perceptions and mass media in the biotechnology controversy. *Int. J. Public Opin. Res.* **2005**, *17*, 5–22. [CrossRef]
2. Kucharavy, D.; De Guio, R. Technological forecasting and assessment of barriers for emerging technologies. In Proceedings of the 17th International Conference on Management of Technology (IAMOT 2008), Dubai, United Arab Emirates, 6–10 April 2008; p. 14-2.
3. Williams, L.; Anthi, E.; Burnap, P. A scalable and automated framework for tracking the likely adoption of emerging technologies. *Information* **2024**, *15*, 237. [CrossRef]
4. Haque, M.U.; Dharmadasa, I.; Sworna, Z.T.; Rajapakse, R.N.; Ahmad, H. " I think this is the most disruptive technology": Exploring Sentiments of ChatGPT Early Adopters using Twitter Data. *arXiv* **2022**, arXiv:2212.05856.
5. Kwarteng, M.A.; Ntsiful, A.; Botchway, R.K.; Pilik, M.; Oplatková, Z.K. Consumer insight on driverless automobile technology adoption via twitter data: A sentiment analytic approach. In Proceedings of the International Working Conference on Transfer and Diffusion of IT, Tiruchirappalli, India, 18–19 December 2020; pp. 463–473.
6. Choudhury, A.; Shamszare, H. Investigating the impact of user trust on the adoption and use of ChatGPT: Survey analysis. *J. Med. Internet Res.* **2023**, *25*, e47184. [CrossRef]
7. OpenAI. ChatGPT: Generative Pre-Trained Transformer, 2023. Available online: https://openai.com/index/chatgpt/ (accessed on 23 August 2024).
8. De Angelis, L.; Baglivo, F.; Arzilli, G.; Privitera, G.P.; Ferragina, P.; Tozzi, A.E.; Rizzo, C. ChatGPT and the rise of large language models: The new AI-driven infodemic threat in public health. *Front. Public Health* **2023**, *11*, 1166120. [CrossRef] [PubMed]
9. Wang, F.Y.; Miao, Q.; Li, X.; Wang, X.; Lin, Y. What does ChatGPT say: The DAO from algorithmic intelligence to linguistic intelligence. *IEEE/CAA J. Autom. Sin.* **2023**, *10*, 575–579. [CrossRef]

10. Davis, F.D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* **1989**, *13*, 319–340. [CrossRef]

11. Venkatesh, V.; Morris, M.G.; Davis, G.B.; Davis, F.D. User acceptance of information technology: Toward a unified view. *MIS Q.* **2003**, *27*, 425–478. [CrossRef]

12. Efuwape, T.O.; Abioye, T.E.; Abdullah, A.K. Text Analytics of Opinion-Poll On Adoption of Digital Collaborative Tools for Academic Planning using Vader-Based Lexicon Sentiment Analysis. *FUDMA J. Sci.* **2022**, *6*, 152–159. [CrossRef]

13. Hizam, S.M.; Ahmed, W.; Akter, H.; Sentosa, I.; Masrek, M.N. Web 3.0 Adoption Behavior: PLS-SEM and Sentiment Analysis. *arXiv* **2022**, arXiv:2209.04900.

14. Mardjo, A.; Choksuchat, C. HyVADRF: Hybrid VADER–Random Forest and GWO for Bitcoin Tweet Sentiment Analysis. *IEEE Access* **2022**, *10*, 101889–101897. [CrossRef]

15. Caviggioli, F.; Lamberti, L.; Landoni, P.; Meola, P. Technology adoption news and corporate reputation: Sentiment analysis about the introduction of Bitcoin. *J. Prod. Brand Manag.* **2020**, *29*, 877–897. [CrossRef]

16. Lee, H.; Kang, P. Identifying core topics in technology and innovation management studies: A topic model approach. *J. Technol. Transf.* **2018**, *43*, 1291–1317. [CrossRef]

17. Cobelli, N.; Blasi, S. Combining topic modeling and bibliometric analysis to understand the evolution of technological innovation adoption in the healthcare industry. *Eur. J. Innov. Manag.* **2024**, *27*, 127–149. [CrossRef]

18. Ansari, K. 500k ChatGPT-Related Tweets Jan-Mar 2023. 2023. Available online: https://www.kaggle.com/datasets/khalidryder777/500k-chatgpt-tweets-jan-mar-2023 (accessed on 16 January 2024).

19. Febriansyah, D. 500k ChatGPT-Related Tweets Jan–Mar 2023. 2023. Available online: https://www.kaggle.com/datasets/danufls/chatgpt-tweets (accessed on 16 January 2024).

20. Domingo, E. ChatGPT 1000 Daily Tweets. 2023. Available online: https://www.kaggle.com/datasets/edomingo/chatgpt-1000-daily-tweets (accessed on 16 January 2024).

21. Anil. Tweets-About-Chatgpt-March-2023. 2023. Available online: https://www.kaggle.com/datasets/sanlian/tweets-about-chatgpt-march-2023 (accessed on 16 January 2024).

22. Banachewicz, K. GPT4-the Tweets. 2023. Available online: https://www.kaggle.com/datasets/konradb/gpt4-the-tweets (accessed on 16 January 2024).

23. Tariq, M. ChatGPT Twitter Dataset. 2023. Available online: https://www.kaggle.com/datasets/tariqsays/chatgpt-twitter-dataset (accessed on 16 January 2024).

24. Bhattacharjee, M. Tweets on ChatGPT-ChatGPT. 2023. Available online: https://www.kaggle.com/datasets/manishabhatt22/tweets-onchatgpt-chatgpt (accessed on 16 January 2024).

25. Pham, M. ChatGPT Tweets First Month of Launch. 2023. Available online: https://www.kaggle.com/datasets/pcminh0505/chatgpt-twitter?resource=download (accessed on 16 January 2024).

26. Rustam, F. ChatGPT Tweets. 2023. Available online: https://www.kaggle.com/datasets/furqanrustam118/chatgpt-tweets (accessed on 16 January 2024).

27. Nutrientaxis. The ChatGPT Tweets Dataset. 2023. Available online: https://www.kaggle.com/datasets/nutrientaxis/the-chatgpt-tweets-dataset/data (accessed on 16 January 2024).

28. Williams, L.; Arribas-Ayllon, M.; Artemiou, A.; Spasić, I. Comparing the utility of different classification schemes for emotive language analysis. *J. Classif.* **2019**, *36*, 619–648. [CrossRef]

29. Ekman, P. Universals and cultural differences in facial expressions of emotion. In *Nebraska Symposium on Motivation*; University of Nebraska Press: Lincoln, NE, USA, 1971.

30. Plutchik, R. A general psychoevolutionary theory of emotion. In *Emotion: Theory, Research, and Experience*; Academic Press: Cambridge, MA, USA, 1980; Volume 1.

31. Laros, F.J.; Steenkamp, J.B.E. Emotions in consumer behavior: A hierarchical approach. *J. Bus. Res.* **2005**, *58*, 1437–1445. [CrossRef]

32. Shaver, P.; Schwartz, J.; Kirson, D.; O'connor, C. Emotion knowledge: Further exploration of a prototype approach. *J. Personal. Soc. Psychol.* **1987**, *52*, 1061. [CrossRef]

33. Storm, C.; Storm, T. A taxonomic study of the vocabulary of emotions. *J. Personal. Soc. Psychol.* **1987**, *53*, 805. [CrossRef]

34. Williams, L.; Anthi, E.; Burnap, P. Comparing hierarchical approaches to enhance supervised emotive text classification. *Big Data Cogn. Comput.* **2024**, *8*, 38. [CrossRef]

35. Valitutti, A.; Strapparava, C.; Stock, O. Developing affective lexical resources. *Psychol. J.* **2004**, *2*, 61–83.

36. Miller, G.A. WordNet: A lexical database for English. *Commun. ACM* **1995**, *38*, 39–41. [CrossRef]

37. Balahur, A.; Steinberger, R.; Kabadjov, M.; Zavarella, V.; van der Goot, E.; Halkia, M.; Pouliquen, B.; Belyaeva, J. Sentiment Analysis in the News. In Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10). European Language Resources Association (ELRA), Valletta, Malta, 17–23 May 2010.

38. Strapparava, C.; Mihalcea, R. Learning to identify emotions in text. In Proceedings of the 2008 ACM Symposium on Applied Computing, Fortaleza, Ceará, Brazil, 16–20 March 2008; pp. 1556–1560.

39. Kang, H.J.; Kim, C.; Kang, K. Analysis of the trends in biochemical research using Latent Dirichlet Allocation (LDA). *Processes* **2019**, *7*, 379. [CrossRef]

40. Williams, L.; Anthi, E.; Arman, L.; Burnap, P. Topic modelling: Going beyond token outputs. *Big Data Cogn. Comput.* **2024**, *8*, 44. [CrossRef]

41. Curiskis, S.A.; Drake, B.; Osborn, T.R.; Kennedy, P.J. An evaluation of document clustering and topic modelling in two online social networks: Twitter and Reddit. *Inf. Process. Manag.* **2020**, *57*, 102034. [CrossRef]

42. Chinnov, A.; Kerschke, P.; Meske, C.; Stieglitz, S.; Trautmann, H. An overview of topic discovery in Twitter communication through social media analytics. In Proceedings of the 20th Americas Conference on Information Systems (AMCIS '15), Puerto Rico, PR, USA, 13–16 August 2015.

43. Spasic, I.; Button, K. Patient Triage by Topic Modeling of Referral Letters: Feasibility Study. *JMIR Med. Inform.* **2020**, *8*, e21252. [CrossRef]

44. Kotuwage, K.; Siyambalapitiya, R. Effects of the COVID-19 pandemic on anxiety disorders: A topic modelling study. *i-Manag. J. Artif. Intell. Mach. Learn. (JAIM)* **2024**, *2*, 18. [CrossRef]

45. Scikit-Learn 0.24.1 Documentation—Latent Dirichlet Allocation. Available online: https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.LatentDirichletAllocation.html (accessed on 5 March 2021).

46. Chang, J.; Gerrish, S.; Wang, C.; Boyd-Graber, J.L.; Blei, D.M. Reading tea leaves: How humans interpret topic models. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 7–10 December 2009; pp. 288–296.

47. Moore, G. *Crossing the Chasm: Marketing and Selling High-Tech Products to Mainstream Customers*; Harper Business Essentials; HarperBusiness: New York, NY, USA, 1991.

48. Rogers, E.M, *Diffusion of Innovations*; Simon and Schuster: New York, NY, USA, 2003; Volume 576.