International Journal of Social Media and Online Communities
Volume 15 • Issue 1

"From Where Should We Work?": Analyzing American Twitter Sentiment of Work Arrangement Preferences

Bryan Ji Yang Leong, State University of New York at Plattsburgh, USA*
https://orcid.org/0009-0003-9020-978X

Zareef Mohammed, Texas State University, USA

ABSTRACT

The COVID-19 pandemic forced organizations to re-evaluate the ways in which they conduct business—specifically, how they allow employees to work. Changes to organizational culture and the demands of today's workforce have brought about a revolution in the work arrangement universe—and a possible death to the traditional 'full-time in office' model. Through the analysis of Twitter sentiment regarding three types of work arrangement preferences, this study observes the preference for flexibility in work modes—a combination working from home + office (hybrid work). Random forest models were able to predict a worker's specific preference to an accuracy of ~68%. Such findings can help businesses continue to sharpen their work arrangement policies that take into consideration trends among the American working populace to ultimately build an effective employment model that will help their organization thrive.

KEYWORDS:

Work Arrangements, Social Media Analytics, Hybrid Work, Work From Home, Work From Office, Twitter, Predictive Models, Topic Modelling, Machine Learning

INTRODUCTION

In recent years, the COVID-19 pandemic has forced organizations and institutions to develop newer work practices heavily involving the use of information systems and the internet. While organizations were constantly changing to adapt to new processes, technology, and even public-health regulations, many of these changes are complex, slow, and mostly ineffective (Howe et al., 2021; Jacobs et al., 2013). Prior to the COVID-19 pandemic, organizations engaged in traditional business practices while skirting around issues/challenges that ultimately created a negative workforce experience. These negative workforce experiences are typically the result of inefficient technologies, archaic ways of working, and poor organizational culture (Bordeaux & Lewis, 2021). As the pandemic progressed,

DOI: 10.4018/IJSMOC.330754
*Corresponding Author

many organizations began to implode from the weight and rate of mass attrition—both forced and voluntary. De Smet et al. (2021) referred to this phenomenon as “the Great Resignation/Reflection,” where workers were leaving their jobs at unprecedented rates and moving to new ones. Similar to a “renter’s market” in the housing industry, the United States (U.S.) was suddenly faced with an “employee’s market,” where employers were now at the mercy of employees who were reflecting on their senses of purpose and meaning at the workplace and opted to look outside their current employment situations to fulfill them (Dhingra et al., 2021).

Deloitte Consulting, a global management consulting firm, described the workforce experience as “the sum of a human’s lived experience at work and how they feel about their organization” (Bordeux & Lewis, 2021). For organizations to retain workers and safeguard human life during the pandemic, many employees received flexible work arrangements. Business meetings shifted to using teleconferencing applications such as Zoom and Microsoft Teams, and employees were able to work from remote locations without needing to physically be at the office. Spreitzer et al. (2017) classified these flexible work arrangements into three dimensions: (1) on-site workers with flexibility in their scheduling, (2) remote workers with a fixed schedule, and (3) flexibility in employment relationships. These changes to the organizational workforce have had lingering effects even after the critical stages of the pandemic. The COVID-19 pandemic was most threatening before the development of the vaccines and boosters, but after the proliferation of vaccines and access to them, organizations—and society in general—began to engage in several pre-pandemic practices, such as reducing the limits of social distancing and mask mandates. The flexible work arrangements offered during the critical stages of the pandemic created a paradigm shift in the lives of many employees, leading them to question the effectiveness, efficiency, and necessity of fixed schedules and permanent work locations.

The flexible working arrangement is a multi-faceted phenomenon with several considerations. As indicated by Spreitzer et al. (2017), one of these considerations pertains to the location at which employees work, which shall become the focus of this study. Specifically, this paper focuses on workers’ preferences of where they would like to work. We apply a data-driven approach to this study by using social-media data (specifically Twitter[1] data) to better understand worker preferences as influenced by the COVID-19 pandemic. Our research is focused on the U.S. population—rather than a global audience—given the possibility of socio-economic and cultural differences that may influence worker preferences (which is outside the scope of this study). Our study is divided in two phases: Phase I pertains to understanding workers’ preferences and perceptions about the work locations, and Phase II aims to develop a suitable predictive model that could be used to determine what category a worker may fall into regarding work preferences. The study is outlined as follows: the next section discusses the research methodology of this study. This is then followed by data analysis and results for Phase I. Consequently, Phase II predictive models and results are then reported. Finally, the paper provides a discussion and conclusion of the study’s key findings.

RESEARCH METHODOLOGY

This study is divided into two phases. Phase I seeks to better understand individuals’ perceptions about working arrangements in the current era of COVID-19 (after the critical stages, where vaccines abound and there is a relaxation of social-distancing rules), while Phase II seeks to develop predictive models to help determine the workers’ work-arrangement preferences. Both phases of this study utilize social-media data for analysis—namely Twitter—and take a data-driven approach to attain insights regarding the phenomenon of work arrangements in the United States.

Social-media analytics has been advocated by researchers as an alternative measure to better understand individuals’ perceptions, even to the point that social-media data may provide enhanced explanations to current social phenomena (Batrinca & Treleaven, 2015; Culotta & Cutler, 2016; Kaplan & Haenlein, 2010; Wamba et al., 2016)[2]. Twitter data have culminated in robust research outcomes across several issues, including consumer insights (Chamlertwat et al., 2012) and perceptions related to marijuana (Cavazos-Rehg et al., 2015). In a similar manner, we believe data derived from tweets can provide unique insight into individuals' work-arrangement preferences.

In this study, we consider three types of working arrangements: (1) work from office (WFO), (2) work from home (WFH), and (3) hybrid (HYB), which is a combinative spectrum between WFO and WFH approaches. Despite the salience associated with the WFO approach in most organizations, WFH approaches have often been studied even prior to the COVID-19 pandemic and are often related to increased satisfaction (Ferguson et al., 2012). Due to the COVID-19 pandemic, many organizations were forced to engage quickly in WFH policies and strategies. When vaccinations were rolled out and less fatal variations of the COVID-19 virus emerged, some organizations reinstituted the WFO approach. For instance, billionaire mogul Elon Musk reportedly "forced" Tesla employees to return to work (Bursztynsky, 2022). However, many organizations retained flexible working arrangements, including the hybrid model (Gratton, 2021). While there may be variations to these approaches, the purpose of this study is to explore the general perceptions of these work arrangements. Hence, we focus specifically on these three work arrangements (WFO, WFH, HYB) from a broader perspective.

We developed several research questions for both Phase I and Phase II. These research questions guided the development of our study using a data-driven, exploratory research approach. The research questions for both phases are presented in Table 1. In the next subsection, we discuss the data collection approach using Twitter.

DATA COLLECTION

Data was collected twice from Twitter with the keywords "wfh" (work from home) "wfo" (work from office), and "hyb" (hybrid work). Initial data collection occurred in April and May of 2022 (30-day collection between the two months), while the second data collection occurred in June 2022, following Elon Musk's edict to return to offices. The reason for the two data-collection samples was to capture tweets (and opinions) before and after Elon Musk's edict to analyze any differences with the sentiments and results in general. Given his prominence as a business leader, it was expected that this announcement would spark worldwide chatter about work arrangements. Since this study is focused on Twitter sentiment across the United States, the location field was scrutinized. Non-U.S. entries were manually filtered from the data, generating datasets of tweets, with the number of observations seen in Table 2. The data were then cleaned of irrelevant tweets (not relating to work arrangements or tweets such as job advertisements) where possible, as well as eliminating stop words, whether the tweet was a job advertisement, and unnecessary (website) links and emoticons. A predictor, "Arrangement," was added to each with its relevant label—WFH, WFO, and HYB—so it could be used as a target variable.

Once the data were collected, we used several data-analysis approaches to derive our results. First, sentiment analysis was conducted on the datasets using various algorithms. Data analysis was conducted using R, and thus, the sentiment-analysis algorithms of R were used for sentiment scores on the three datasets. Each document (tweet) was analyzed through the lens of the Syuzhet, AFINN,

Table 1. Research questions

Phase	Research Questions
Phase I—Exploration of Work-Arrangement Perceptions	1. What are the general sentiments about various work conditions in the U.S.? 2. Are these sentiments similar among average persons? 3. How did sentiments change following Elon Musk's WFO edict in June 2022? 4. Does sentiment about work arrangements vary across different states in the U.S.? 5. What are the key topics being discussed about work arrangements?
Phase II—Predictive Model of Working-Arrangement Preferences	1. How accurately can we use various characteristics of a tweet to predict an individual's work preference?

Table 2. Number of observations per work-arrangement subset

Work-Arrangement Dataset	Number of Tweets
Work from Home (WFH)	3,139
Work from Office (WFO)	1,938
Hybrid Work (HYB)	1,376

and Bing sentiment algorithms, which provided a scaled score for positive and negative sentiments based on each word of a tweet. We also applied the SentimentR algorithm, which facilitates the context of a sentence, therefore providing a better measure of user sentiment. These were used in Phase I of our analysis, among other methods such as topic modeling and visualizations, to answer the research questions of this phase. For Phase II, we used the sentiments (specifically those derived from AFINN and SentimentR) as features of the predictive models, along with the number of likes, verified status, number of followers, and location. The next section details the data-analysis process and the associated results.

PHASE I: DATA ANALYSIS AND RESULTS

In this section, we present the data analysis and results derived from the data collection from Phase I of this study.

Sentiments Concerning Work Arrangements

The first research question for Phase I was: what are general sentiments about the various work arrangements in the U.S.? To answer this research question, sentiment analysis using the algorithms in the R language was used. Naldi and Petroni (2023) conducted a review on the sentiment-analysis algorithms in R in which they concluded that SentimentR was the only algorithm that accounted for negators. Other algorithms, such as those found in the Syuzhet package (i.e., Syuzhet, Bing, AFINN, and NRC), would consider the emotional valence of a word as its own separate score (for instance, “happy” would be considered positive irrespective of whether it was negated by “not” before it). Hence, in this study, SentimentR holds the most significance in determining the overall perceptions or sentiments individuals have about working arrangements in the U.S. However, we also included AFINN, Bing, and Syuzhet as comparisons, as well as the value each of these algorithms brings. While these algorithms are limited by their considerations of negators, they do not completely ignore negators. After all, the sentiment scores for these three algorithms are a summation of the positive and negative words in a document. Therefore, a sentence with a positive word and a negative word may result in a neutral score (Naldi & Petroni, 2023). Most documents tend to contain several positive and negative words, each of which have various weights depending on the algorithm. This in turn provides an overall sentiment insight. Table 3 presents the average sentiment scores of each of the working arrangements per algorithm.[3]

The net positive averages of sentiment per work arrangement indicate that the words associated with various tweets are generally positive. However, HYB work arrangements score the highest among the three work arrangements in all sentiment-analysis algorithms. Surprisingly, WFO is higher than WFH, indicating that individuals may prefer working from the office as opposed to working from home—while still enjoying some flexibility to their work arrangements.

Sentiments About Work Arrangements for the Average Person

Following the first research question, the second research question is: are these sentiments similar among average persons? For this research question, we excluded any verified tweets. The verified

Table 3. Sentiment analysis of work arrangements

	WFH	HYB	WFO
Syuzhet	0.5931	1.5882	0.9314
Bing	0.5021	1.6569	1.1425
AFINN	0.9997	1.8813	1.1483
SentimentR	0.0699	0.2001	0.0957

status on Twitter indicates that the username operates on an account of public interest—essentially, famous users.[4] Hence, we refer to nonverified users as "average," as they are not prominent users but rather encapsulate what can be considered the average Twitter user. The number of tweets remaining after removing famous users is displayed in Table 4, part a. The results, presented in Table 4, part b, show similar results to those from research question 1, where individuals seem to prefer work from office, with some degree of flexibility (i.e., hybrid approaches) more than work from home.

Sentiments About Work Arrangements Across States

The third research question for Phase I is as follows: does the sentiment about work arrangements change across various states in the U.S.?

To best approach this, we manually cleaned and transformed the location predictor variable to identify the U.S. state where the tweet originated from. The dataset was loaded into Tableau and visualized onto a map of America based on the average sentiment score for each state. The SentimentR algorithm was used for this analysis, as it provided the most distinguishable separation of classes among the other sentiment metrics. Figure 1 depicts the sentiments by state for the WFH work arrangement, while Figures 2 and 3 depict the sentiments by state for HYB and WFO, respectively.

As seen in Figure 1, the darker shade of green represents a more positive average sentiment, while the red hues represent negative sentiments[5]. The average sentiments about WFH arrangements are generally positive across the U.S., and specific states such as Montana, Nevada, and Arkansas have the strongest positive sentiments. Most other states show relatively strong affinity toward WFH as well. However, New Mexico and Mississippi tend to have a negative average sentiment, while South

Table 4a. Number of observations per work-arrangement subset

Work-Arrangement Dataset	Number of Total Tweets (From Verified and Unverified Users)	Number of Tweets From Average (Not Verified) Users	Difference
Work from Home (WFH)	3,139	2,981	-158
Work from Office (WFO)	1,938	1,758	-180
Hybrid Work (HYB)	1,376	1,249	-127

b. Sentiment analysis of work arrangements of the average person

	WFH	HYB	WFO
Syuzhet	0.5932	1.5553	0.9125
Bing	0.5073	1.6315	1.4049
AFINN	0.9883	1.7934	1.0794
SentimentR	0.0701	0.2021	0.0955

Figure 1. WFH sentiment analysis across U.S. states

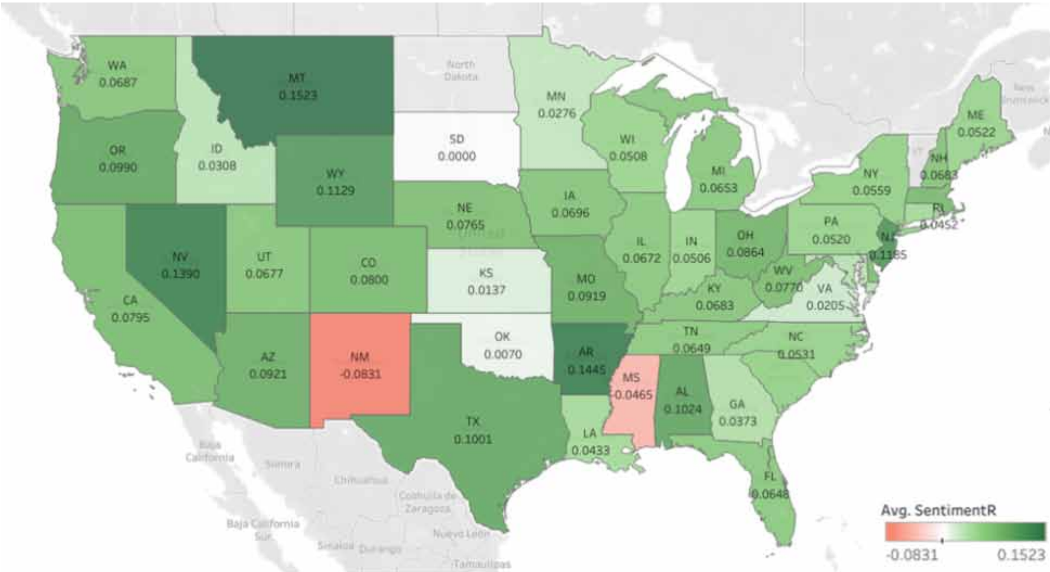
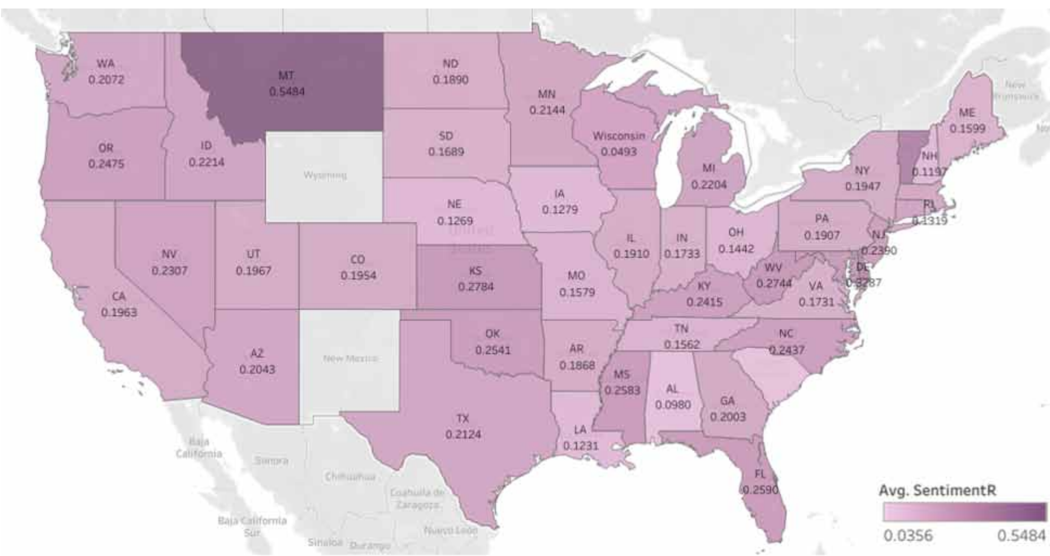


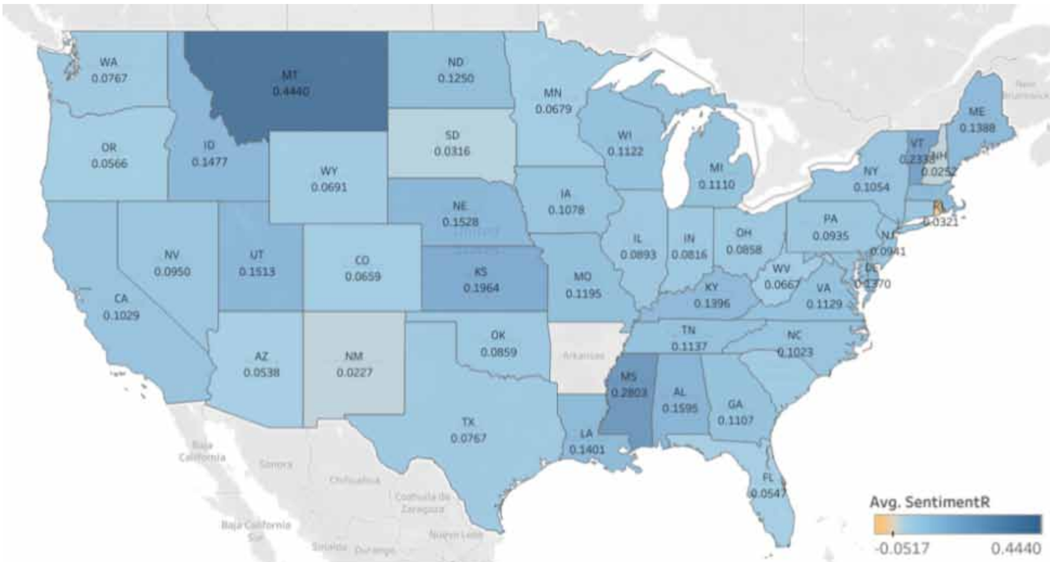
Figure 2. HYB sentiment analysis across U.S. states



Dakota, Kansas, and Oklahoma appear to be neutral (with an average sentiment score of or close to 0) about the WFH arrangement.

When analyzing HYB work-arrangement preference across states, the average sentiment is positive in all states (with no negative average sentiment). Both Montana and Vermont have the highest positive average sentiments, while other states such as Delaware, West Virginia, and Kansas also show a better-than-average positive average SentimentR score, indicating preference for hybrid work arrangements in these areas.

Figure 3. WFO sentiment analysis across U.S. states



South Carolina and Alabama appear to be generally more neutral about the HYB work arrangement (with average sentiment scores in the positives, but somewhat closer to 0). This observation is especially fair for Alabama—a state that also tended to have more positive sentiment about the WFH arrangement (seen in the previous analysis).

For the WFO arrangement, the strongest positive preferences (2.3 average score) were Vermont, Mississippi, and Montana. Interestingly, Vermont shows a decent spread of positive sentiment for hybrid and work from office arrangements. Montana records the highest positive sentiment scores across all three types of work arrangements, indicating a strong mix of preferences for all three types of work (home, office, and hybrid). Overall, all states show positive affinity toward working from the office, apart from Rhode Island (-0.0321 average score).

In summary, most states show positive affinity to the three work arrangements. However, it should be noted that the sample of tweets per state is unequal. While this analysis gives a general idea of how individuals of each state may feel toward the three work arrangements, some states constituted a very small sample of tweets. This is a consequence of collecting data from Twitter or any social-media platform where samples are dynamic and vary over time. Unlike traditional data-collection methods (e.g., surveys), researchers are often unable to control all factors in the data-collection procedure.

Change of Sentiments Following Elon Musk's WFO Edict

The fourth research question is: how did sentiments change following Elon Musk's WFO edict? At the end of May 2022, Tesla employees were instructed by CEO Elon Musk to return to the office or face termination (Boyle, 2022; Bursztynsky, 2022; Mac, 2022). This edict by the billionaire tech mogul fueled the flames of a continued debate as society moved into a post-pandemic world: should employees now return to work in the office?

Approximately a week after Musk's mandate, we scraped data from Twitter again to allow a reasonable collection of robust conversation about the topic. The tweets were cleaned using the same process as seen earlier for the first research question. Table 5 shows the number of tweets remaining after the dataset was cleaned for irrelevant locations (non-U.S.).

Table 6, parts a, b, and c report the average sentiments before and after Elon's edict (EE) on work arrangements (WFH, HYB, and WFO, respectively). Each table also presents the percentage

Table 5. Number of observations per work-arrangement subset

Work-Arrangement Dataset	Number of Tweets
Work from Home (WFH)	3,578
Work from Office (WFO)	2,209
Hybrid (HYB)	1,953

Table 6a. Impact of Elon's edict (EE) on WFH sentiment

WFH Sentiment	Before EE	After EE	Percentage Change
Syuzhet	0.5931	0.5597	-6%
Bing	0.5021	0.5240	+4%
AFINN	0.9997	0.7868	-27%
SentimentR	0.0699	0.1365	+49%

b. Impact of Elon's edict (EE) on HYB work sentiment

HYB Sentiment	Before EE	After EE	Percentage Change
Syuzhet	1.5882	1.3908	-14%
Bing	1.6569	1.5166	-9%
AFINN	1.8813	1.3676	-38%
SentimentR	0.2001	0.3745	+47%

c. Impact of Elon's edict (EE) on WFO sentiment

WFO Sentiment	Before EE	After EE	Percentage Change
Syuzhet	0.9314	0.7045	-32%
Bing	1.1425	1.1987	+5%
AFINN	1.1483	0.5708	-101%
SentimentR	0.0957	0.1795	+47%

change of the sentiments before EE and after EE. As is noticeable, the results differ for different sentiment algorithms. For the WFH arrangement, sentiments decreased after Elon's edict based on the Syuzhet and AFINN algorithms. The same is true for HYB as well as WFO. However, Bing saw an increase for both WFH and WFO arrangements, but a small decrease for HYB. These findings are inconsistent for the specific work arrangement (for example—concerning WFH—Syuzhet and AFINN have negative percentage changes, while Bing has a small positive percentage change). It is possible that this represents the drawbacks of these types of algorithms that do not account for the context of a statement, i.e., negatory words (Naldi & Petroni, 2023).

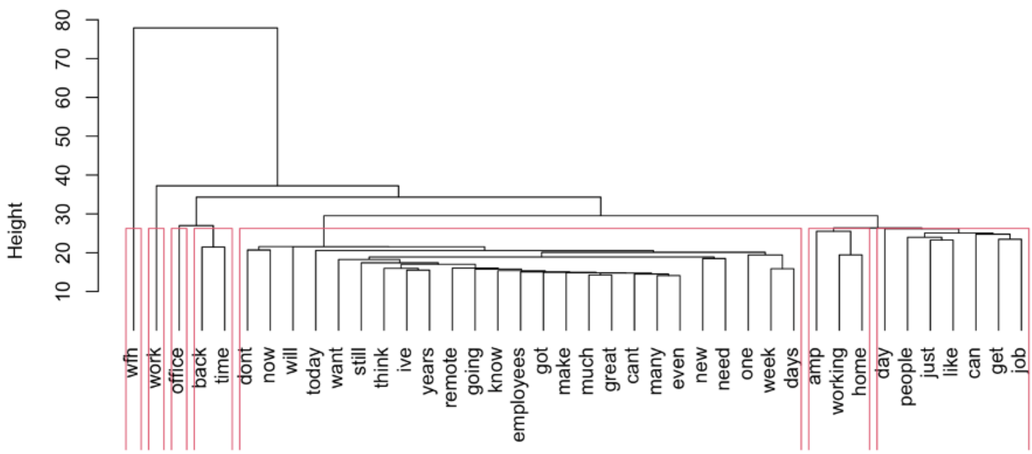
The SentimentR algorithm revealed an increase in the percentage change after Elon's edict for all three work arrangements. These percent changes are also very similar (49%, 47%, and 47% for WFH, HYB, and WFO, respectively). Interestingly, when we look at actual sentiment scores, HYB work arrangements score the highest—whether before or after Elon's edict. These results do not show any significant change in sentiments following Elon's back to office edict, indicating that it may not have changed workers' opinions or feelings.

Key Topics Pertaining to Work Arrangements

The fifth and final research question for Phase I is: what are some key topics being discussed about work arrangements? We used two unsupervised machine-learning approaches to address this research question: hierarchical clustering and Latent Dirichlet Allocation (LDA). These modeling approaches were conducted for each dataset of tweets (WFH, WFO, and HYB). Although both approaches were used, hierarchical clustering was used primarily to determine the number of topics to include in the LDA; the latter was more intuitive to the analysis process of topic modeling using Twitter data.

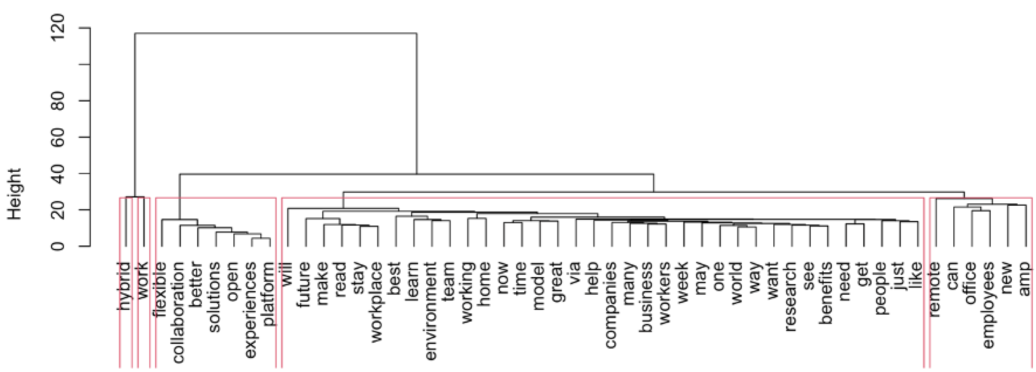
WFH

Figure 4. Dendrograms of hierarchical clustering for work from home arrangement



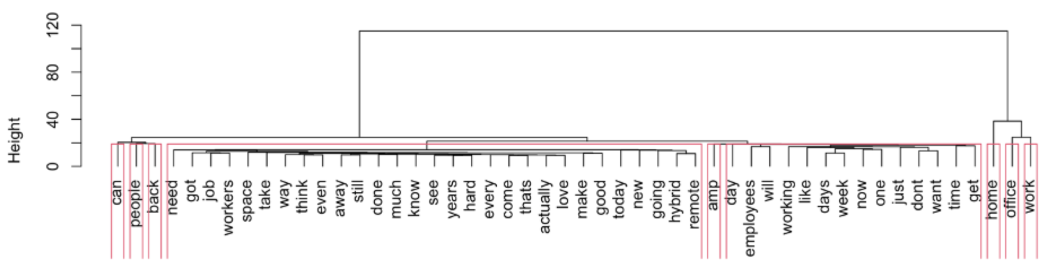
HYB

Figure 5. Dendrograms of hierarchical clustering for hybrid work arrangement



WFO

Figure 6. Dendrograms of hierarchical clustering for work from office arrangement



Figures 4, 5, and 6 display the hierarchical clustering dendrograms for WFH, HYB, and WFO tweets, respectively. First, for WFH, the model identified seven main topics: (1) working from home, (2) work, (3) office, (4) time associated with WFH, (5) time associated with remote work, (6) benefits of working at home, and (7) the job market. Next, the HYB clustering approach identified five main topics: (1) hybrid work arrangements, (2) work, (3) flexibility, (4) benefits of hybrid work arrangements, and (5) innovation of this work-arrangement approach. Finally, the WFO clustering model resulted in eight topics (two of which were discarded, as the word for each associated cluster provides little context): (1) people, (2) return to traditional work arrangements, (3) pros and cons of WFO after different approaches were experienced, (4) time associated with WFO, (5) home, (6) office, and (7) work. Figures 7, 8, and 9 display the results of the LDA models for WFH, HYB and WFO, respectively.

WFH

Figure 7. Results of LDA model for work from home tweets

```
> terms(wfh_topics, 15)
      Topic 1  Topic 2    Topic 3  Topic 4     Topic 5
 [1,] "wfh"    "wfh"      "wfh"    "job"       "work"
 [2,] "office" "can"      "just"   "amp"       "wfh"
 [3,] "time"   "people"   "like"   "working"   "day"
 [4,] "back"   "one"      "dont"   "will"      "home"
 [5,] "now"    "still"    "need"   "new"       "get"
 [6,] "days"   "ive"      "know"   "remote"    "today"
 [7,] "going"  "also"     "great"  "employees" "cant"
 [8,] "years"  "return"   "even"   "make"      "take"
 [9,] "week"   "business" "got"    "jobs"      "every"
[10,] "since"  "pandemic" "good"   "want"      "youre"
[11,] "best"   "see"      "really" "workers"   "getting"
[12,] "lol"    "pay"      "think"  "company"   "feel"
[13,] "covid"  "well"     "life"   "much"      "right"
[14,] "hours"  "less"     "thats"  "doesnt"    "never"
[15,] "first"  "play"     "want"   "companies" "help"
```

HYB

Figure 8. Results of LDA model for hybrid work tweets

```
> terms(hyb_topics, 15)
      Topic 1    Topic 2     Topic 3         Topic 4     Topic 5
 [1,] "work"     "new"       "hybrid"        "office"    "work"
 [2,] "hybrid"   "amp"       "work"          "employees" "hybrid"
 [3,] "working"  "best"      "can"           "will"      "remote"
 [4,] "week"     "business"  "flexible"      "home"      "via"
 [5,] "model"    "one"       "future"        "now"       "may"
 [6,] "job"      "research"  "environment"   "people"    "see"
 [7,] "schedule" "like"      "learn"         "time"      "virtual"
 [8,] "want"     "way"       "collaboration" "get"       "powering"
 [9,] "also"     "workforce" "team"          "just"      "employee"
[10,] "check"    "practices" "workplace"     "great"     "culture"
[11,] "looking"  "going"     "better"        "help"      "event"
[12,] "day"      "company"   "make"          "companies" "take"
[13,] "youre"    "look"      "read"          "workers"   "healthcare"
[14,] "even"     "strategy"  "solutions"     "many"      "report"
[15,] "hiring"   "ways"      "open"          "back"      "experience"
```

WFO

Figure 9. Results of LDA model for work from office tweets

```
> terms(wfo_topics, 15)
      Topic 1   Topic 2  Topic 3    Topic 4  Topic 5
 [1,] "just"    "work"   "office"   "home"   "work"
 [2,] "day"     "office" "work"     "work"   "office"
 [3,] "time"    "amp"    "back"     "office" "home"
 [4,] "working" "now"    "people"   "dont"   "can"
 [5,] "get"     "new"    "want"     "going"  "employees"
 [6,] "like"    "hybrid" "austin"   "remote" "will"
 [7,] "one"     "good"   "workers"  "make"   "space"
 [8,] "need"    "may"    "see"      "know"   "anywhere"
 [9,] "today"   "think"  "actually" "take"   "since"
[10,] "days"    "great"  "company"  "much"   "use"
[11,] "got"     "best"   "next"     "way"    "thats"
[12,] "job"     "team"   "get"      "really" "help"
[13,] "away"    "future" "part"     "also"   "full"
[14,] "even"    "staff"  "via"      "done"   "well"
[15,] "years"   "hard"   "young"    "week"   "told"
```

In general, the LDA model provided more sufficient word associations. This allowed for better interpretation of the possible topics discussed on the three different work arrangements. The results do suggest that each work arrangement heavily centers around time, as it is the first topic identified for each work arrangement. In addition, we see more advocacy of changing traditional work arrangements to include more hybrid or remote work. From the analysis in Table 7, we observe how workers seem to prefer discussions around moving away from a solely traditional WFO setting.

PHASE II: PREDICTIVE MODELS OF WORKING-ARRANGEMENT PREFERENCES

The second phase of this study was to develop predictive models that could determine an individual's working-arrangement preferences using social media data. To accomplish this task, we used work arrangements as the response variable with three classes: wfo, wfh, and hyb. Our predictors included likes, verified status (true or false), number of followers, and two sentiment scores: AFINN and SentimentR. We used two sentiment scores because of the differences in the algorithms, where one is more sensitive to negatory words (SentimentR) and the other (AFINN) provides a score based on the number of positive and negative words in a document. Since both these sentiment-analysis algorithms are different from each other in how they operate, we included them in the model. SentimentR is more context-specific, while AFINN was selected among the various sentiment measures as it is acknowledged to be built to analyze Twitter-related sentiment (Sonkin, 2021).

We developed several classification models to determine which would provide the highest accuracy for predicting work arrangements. Classification-supervised machine-learning models were used since the outcome variable was categorical with three classes. Since the datasets between classes were imbalanced, we performed both undersampling and oversampling on the datasets to adjust for these imbalances. Then we used all three datasets (original, undersampled, and oversampled) in the classification models to determine the best model. Table 8 presents the sample sizes for each dataset and their classes.

Table 7. Analysis of work-arrangements topics based on clustered keywords

Topics	WFH	HYB	WFO
Topic 1	Time associated with WFH	Scheduling/time under HYB work arrangements	Time associated with working from office
Topic 2	The end of WFH and return to WFO	Call to research—inform workforce practice and strategy	Work arrangements innovation—methods like hybrid approaches and their benefits over WFO
Topic 3	Benefits of working from home	Benefits of hybrid approach—collaboration and flexibility	People's wants pertaining to WFO
Topic 4	Shift of employees' expectations to some form of working from home	Time associated with HYB work arrangements	Lack of motivation for WFO arrangements
Topic 5	Struggles of WFH	Culture and empowerment of HYB	Disconnection of effective work with physical space

Table 8. Variations of original, undersampled, and oversampled datasets for supervised models

	HYB	WFH	WFO
Original Dataset	2,244	3,599	1,662
Undersampled	1,500	1,500	1,500
Oversampled	3,228	3,599	3,324

In totality, we developed the following supervised models:

1. Multimodal logistic regression—needed to handle more than two classes as opposed to logistic regression.
2. Artificial neural network with two hidden nodes.
3. Artificial neural network with three hidden nodes.
4. Support vector machine (with linear kernel).
5. Support vector machine (with radial kernel).
6. Naïve Bayes classifier.
7. Classification tree (unpruned, depth unspecified).
8. Classification tree (pruned).
9. Random forest (500 trees).
10. Random forest (1,000 trees).

Some supervised machine-learning models were not used, such as k-nearest neighbors or linear/quadratic discriminant analysis. These models would be unsuitable given that some of the predictors are categorical variables. Table 9 displays the accuracy, sensitivity, and specificity of each model for the original dataset, while Tables 10 and 11 present the results for the undersampled and oversampled datasets, respectively. In general, the highest accuracy we attained was from the random forest model with 500 trees at 68.12%. While this may be considered somewhat low, it is worth remembering that the predictions are based on social-media data—more specifically, from interpretations of individuals' responses via sentiment analysis.

In the next section, we discuss the implications of these findings for both phases of our research, as well as possible improvements that can be made for future research.

DISCUSSION AND CONCLUSION

The COVID-19 pandemic forced organizations to re-evaluate the ways in which they conduct business. While the pandemic is in a healing phase due to the dissemination of the vaccine, less fatal mutations of the virus, and general herd immunity, its aftereffects continue to shape the conversation around work arrangements across the United States—arguably changing it forever. Technology and flexible schedules allow individuals freedom in their work preferences, in both their selected hours and their work locations. However, organization leaders may not see such work arrangements as desirable, with some even calling for a return to the office. Yet demands of the workforce and changes in organizational culture have brought about a revolution in the work-arrangement universe—and the possible death of the traditional office model of work. In this study, we sought to better understand worker preferences as influenced by the COVID-19 pandemic using a data-driven approach that analyzed these sentiments via Twitter data. Furthermore, we developed predictive models to determine a work-arrangement preference category into which a worker would fall based on their tweets. Our research revealed several insights into the U.S. population's stance on work preferences, while providing a possible solution for businesses to better organize their workforce.

Our research yielded several interesting findings in both phases of the study. Results from Phase I suggest that Americans perceive hybrid work arrangements as more positive and, by extension, prefer the hybrid model in comparison to working from home or from the office. Interestingly, our findings do indicate that working from the office has a higher sentiment than working from home. This may be due to worker preferences for some levels of socialization that would otherwise be limited when working from home. When we filtered out "famous" tweets from the Twitter data, the results were similar.

Next, our research revealed that most U.S. states—besides New Mexico and Mississippi—have positive sentiments toward working from home. Mississippi tends to feel more positively about office

Table 9. Original dataset

Supervised Learning Model	Base Dataset							
	Model ID	Accuracy (%)	WFH Class		HYB Class		WFO Class	
			Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
Multinomial logistic regression	mn1	53.52	84.24	30.15	39.74	84.08	00.00	99.90
Artificial neural network (2 hidden nodes)	nn1	50.39	96.65	7.39	11.86	97.17	00.00	100.0
Artificial neural network (3 hidden nodes)	nn2	50.93	95.53	9.96	15.63	96.08	00.00	100.0
Support vector machine (linear kernel)	svm1	54.07	88.61	24.92	34.21	88.47	00.00	100.0
Support vector machine (radial kernel)	svm2	54.14	83.62	38.15	42.63	83.97	0.370	99.71
Naïve Bayes classifier	nB1	51.51	69.11	48.46	51.05	76.18	10.37	92.75
Classification trees (unpruned, depth unspecified)	ct1	56.78	82.68	43.08	48.95	82.66	6.296	97.06
Classification trees (pruned)	ct2	55.15	73.95	53.69	62.63	69.48	00.00	100.0
Random forest (500 trees)	rf500	54.07	69.58	56.00	57.89	73.22	11.85	93.83
Random forest (1,000 trees)	rf1000	54.38	70.51	54.92	57.37	74.42	11.85	93.83

work and hybrid work, indicating that there is a case to be made for these types of work arrangements in the state. All states (except Wyoming and New Mexico, where no tweets were scraped during the study timeframe) show a positive sentiment toward the hybrid work arrangement, particularly Montana and Vermont (these states have the highest positive sentiment score toward hybrid work). Vermont displayed positive sentiment toward both the office and hybrid arrangements of work, but there were no tweets from users of this region associated with the work from home arrangement. Interestingly, sentiment from Montana indicates a strong mix of preference for all three types of arrangements. This could bring many implications for the types of work-arrangement policies that can be implemented there.

After analyzing the sentiments pertaining to each state, we re-collected data following Elon Musk's edict that all employees must return to the office to work. The results did not contain any drastic changes in sentiments, indicating that Elon's edict may not have changed workers' opinions or feelings about the topic. Next, we used a topic modeling approach, to which we found five topics each associated with each work-arrangement type (see Table 7). In general, the topics suggest that hybrid and work from home approaches are seen more positively, while the topics associated with work from office are perceived more negatively.

Phase II of our study focused on developing predictive (supervised) machine learning models to determine the work-arrangement preference category into which a tweeter would fall based on the

Table 10. Undersampled dataset

Supervised Learning Model	Undersampled Dataset							
	Model ID	Accuracy (%)	WFH Class		HYB Class		WFO Class	
			Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
Multinomial logistic regression	mn2	43.33	64.26	50.57	53.50	72.35	11.86	92.23
Artificial neural network (2 hidden nodes)	nn3	35.78	95.19	9.52	9.55	98.46	5.09	97.03
Artificial neural network (3 hidden nodes)	nn4	34.89	95.19	9.20	9.87	98.64	2.03	95.87
Support vector machine (linear kernel)	svm3	44.22	54.64	66.67	49.68	75.26	28.14	74.55
Support vector machine (radial kernel)	svm4	47.00	40.55	80.13	58.60	68.94	41.02	71.24
Naïve Bayes classifier	nB2	47.89	50.86	69.79	66.56	65.19	25.09	86.61
Classification trees (unpruned, depth unspecified)	ct3	49.22	42.27	76.03	50.96	77.65	54.24	70.25
Classification trees (pruned)	ct4	49.56	44.33	81.28	75.48	56.66	27.12	85.79
Random forest (500 trees)	rf500_U	60.89	48.11	84.73	68.79	75.26	65.08	81.16
Random forest (1,000 trees)	rf1000_U	59.78	47.77	83.09	66.56	75.43	64.41	80.99

content of their tweets. Of the 10 models created, the best performing model is the random forest with 500 trees, with an accuracy of 68.12% using an oversampled dataset (see Tables 9–11 for results on each model's accuracy and the sensitivity and specificity for each class). Each model was developed using the same features (or predictors) from Twitter data, which included the number of likes, verified status, number of followers, and two sentiment scores from the AFINN and SentimentR algorithms. The higher accuracy of the oversampled dataset indicates the importance of class equality.

Contributions

Human resources remain one of the most important elements of a successful organization, and research has suggested that flexible work arrangements can impact various outcomes, including job satisfaction (Niebuhr et al., 2022). Our study contributes to better understanding work preferences following the COVID-19 pandemic through several key insights. First, the results provide general insight into an individual's overall work-arrangement preferences. These insights can help organizations better understand their workforce and develop solutions to balance employee desires and business productivity.

Second, our study demonstrates the utility of social-media data as a solution for business decisions. In this case, social-media data (particularly Twitter data) were used in both analyzing worker perspectives and predicting worker affinity toward certain work arrangements. The results of this study can be easily replicated by organizations from various industries to better understand their specific workforces. Specifically, companies can analyze the social-media data (whether mainstream

Table 11. Oversampled dataset

Supervised Learning Model	Oversampled Dataset							
	Model ID	Accuracy (%)	WFH Class		HYB Class		WFO Class	
			Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
Multinomial logistic regression	mn3	44.35	72.15	40.46	47.24	75.17	0.075	97.51
Artificial neural network (2 hidden nodes)	nn5	42.87	25.38	85.35	78.99	54.05	29.43	76.63
Artificial neural network (3 hidden nodes)	nn6	46.64	37.61	81.70	79.38	56.00	26.04	83.39
Support vector machine (linear kernel)	svm5	43.00	77.02	33.27	44.29	77.10	00.00	100.0
Support vector machine (radial kernel)	svm6	48.29	58.45	63.95	66.93	65.94	18.13	91.67
Naïve Bayes classifier	nB3	45.88	53.27	64.33	60.24	68.46	23.18	85.24
Classification trees (unpruned, depth unspecified)	ct5	58.47	64.08	74.59	71.65	78.02	39.07	84.64
Classification trees (pruned)	ct6	49.24	57.38	71.91	66.34	68.20	22.99	83.61
Random forest (500 trees)	rf500_O	68.12	67.43	81.98	83.46	82.72	54.39	87.30
Random forest (1,000 trees)	rf1000_O	67.88	66.67	82.93	84.06	81.29	54.02	87.55

social media, corporate-engagement surveys, or even—where appropriate—organizational social media such as Slack, Teams, etc.) from their employees to better understand their perspectives. This will help organizations to tailor work-arrangement solutions for current employees while leveraging similar machine-learning models to determine the preferences of future hires. Our results indicate that the best model for such a task is the random forest with 500 trees. While newer data may change the performances of these models, the random forest model can be a starting point; researchers can build, test, and sharpen other models where needed to drive greater predictive accuracy.

Limitations and Future Research

Our study was subject to some limitations, one of which is that only Twitter data were analyzed. As a remedy, future research should leverage other social-media platforms and dimensions (such as demographics), as more insight may be drawn from them. Concerning demographics, research by Suarez (2022) highlights aspects of socio-economic differences in remote work arrangements that can be used as a springboard in this avenue.

Studies can focus on improving Phase II models with alternative or additional social-media data. For instance, data derived from LinkedIn may provide even better insights, as that social-media platform is a dedicated professional network. Additionally, the accuracy of our best performing model was 68.12%, a possible consequence of using specific types of algorithms to derive sentiment scores. Similarly, the limited number of features used to build each model could also cause this. This limitation can be addressed in future studies by utilizing different features, better algorithms, cross-validation approaches, and—as discussed in the first limitation—a combination of data from alternative social-media platforms. Such variety will likely help studies attain more robust insights and predictive models with higher accuracy.

Furthermore, one of the features, "followers," could have contained ghost profiles, which could have falsely increased the number of a user's followers. A ghost profile is someone that is inactive but has not deactivated their account. Research has revealed that there are a variety of users in online communities, one of which is considered as "retrieve information users." These are passive users only involved in gaining information and content (de Valck et al., 2009). To summarize, it is difficult to assess whether a user (1) is inactive because they no longer have an interest in social media, (2) created fake accounts just to increase the followers for another account, or (3) simply browses social media on occasion. Given the complexity of this issue and the fact that its conceptual value was outside the scope of this study, we suggest future studies can incorporate this into their models. In any case, we do not believe this has significantly affected our models, given that it relates to only one of our features.

Some might note that the Phase I finding suggesting that workers prefer working from the office more than working from home is counterintuitive to the anecdotal sentiment of the day. Future studies should consider circumstantial data not captured in this study. For example, the family situation of the individual worker (i.e., whether they are a carer for children/elderly) could provide an interesting insight on their preference and add more depth to the discussion between affinity toward working from home or working from the office. Identifying such specific work-life features could provide insight to individual organizations and help them identify a solid starting point on which they are able to build an effective employment model that will help their employees and business thrive.

REFERENCES

Alexander, A., De Smet, A., Langstaff, M., & Ravid, D. (2021, November). *What employees are saying about the future of work*. McKinsey & Co. https://www2.mvcc.edu/shn/pdf/presentations/2021-11/what-employees-are-saying-about-the-future-of-remote-work.pdf

Batrinca, B., & Treleaven, P. C. (2015). Social media analytics: A survey of techniques, tools and platforms. *AI & Society*, *30*(1), 89–116. doi:10.1007/s00146-014-0549-4

Bordeaux, C., & Lewis, S. (2021, September 23). *Designing the workforce experience with the human at the center.* Deloitte. https://www2.deloitte.com/us/en/blog/human-capital-blog/2021/human-centered-workforce-experience.html

Boyle, M. (2022, June 1). *Elon Musk's demand staffers stop 'phoning it in' may cost him talent*. Bloomberg. https://www.bloomberg.com/news/articles/2022-06-01/musk-s-return-to-office-mandate-flies-in-the-face-of-new-reality

Bursztynsky, J. (2022, June 1). *Elon Musk tells Tesla workers to return to the office full time or resign.* CNBC. https://www.cnbc.com/2022/06/01/elon-musk-reportedly-tells-tesla-workers-to-be-in-office-full-time-or-resign.html

Cavazos-Rehg, P. A., Krauss, M., Fisher, S. L., Salyer, P., Grucza, R. A., & Bierut, L. J. (2015). Twitter chatter about marijuana. *The Journal of Adolescent Health*, *56*(2), 139–145. doi:10.1016/j.jadohealth.2014.10.270 PMID:25620299

Chamlertwat, W., Bhattarakosol, P., Rungkasiri, T., & Haruechaiyasak, C. (2012). Discovering consumer insight from Twitter via sentiment analysis. *Journal of Universal Computer Science*, *18*(8), 973–992.

Culotta, A., & Cutler, J. (2016). Mining brand perceptions from Twitter social networks. *Marketing Science*, *35*(3), 343–362. doi:10.1287/mksc.2015.0968

De Smet, A., Dowling, B., Mugayar-Baldocchi, M., & Schaninger, B. (2021, September 8). *'Great attrition' or 'great attraction'? The choice is yours.* McKinsey Quarterly. https://www.mckinsey.com/business-functions/people-and-organizational-performance/our-insights/great-attrition-or-great-attraction-the-choice-is-yours

De Valck, K., Van Bruggen, G. H., & Wierenga, B. (2009). Virtual communities: A marketing perspective. *Decision Support Systems*, *47*(3), 185–203. doi:10.1016/j.dss.2009.02.008

Dhingra, N., Samo, A., Schaninger, B., & Schrimper, M. (2021, April 5). *Help your employees find purpose—or watch them leave*. McKinsey & Co. https://www.mckinsey.com/business-functions/people-and-organizational-performance/our-insights/help-your-employees-find-purpose-or-watch-them-leave

Ferguson, M., Carlson, D., Zivnuska, S., & Whitten, D. (2012). Support at work and home: The path to satisfaction through balance. *Journal of Vocational Behavior*, *80*(2), 299–307. doi:10.1016/j.jvb.2012.01.001

Gratton, L. (2021). *How to do hybrid right.* Harvard Business Review. https://hbr.org/2021/05/how-to-do-hybrid-right

Howe, D. C., Chauhan, R. S., Soderberg, A. T., & Buckley, M. R. (2021). Paradigm shifts caused by the COVID-10 pandemic. *Organizational Dynamics*, *50*(4), 100804. doi:10.1016/j.orgdyn.2020.100804 PMID:33191959

Jacobs, G., van Witteloostuijn, A., & Christe-Zeyse, J. (2013). A theoretical framework of organizational change. *Journal of Organizational Change Management*, *26*(5), 772–792. doi:10.1108/JOCM-09-2012-0137

Kaplan, A., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons*, *53*(1), 59–68. doi:10.1016/j.bushor.2009.09.003

Mac, R. (2022, June 1). *Elon Musk to workers: Spend 40 hours in the office, or else*. The New York Times. https://www.nytimes.com/2022/06/01/technology/elon-musk-tesla-spacex-office.html

Naldi, M., & Petroni, S. (2023). A testset-based method to analyze the negation-detection performance of lexicon-based sentiment analysis tools. *Computers*, *12*(18), 1–17.

Niebuhr, F., Borle, P., Borner-Zobel, F., & Voelter-Mahlknecht, S. (2022). Healthy and happy working from home? Effects of working from home on employee health and job satisfaction. *International Journal of Environmental Research and Public Health*, *19*(3), 1122. doi:10.3390/ijerph19031122 PMID:35162145

APPENDIX A

In this appendix, researchers reran the sentiment analysis by undersampling the datasets (where all three datasets have the same number of observations). While the sentiment scores changed—as is normal whenever the sample size changes—the inferences are the same as in the section entitled "Sentiments Concerning Work Arrangements" above.

Table A1. Sentiment analysis of work arrangements (with undersampled dataset)

	WFH	(Std. Deviation)	HYB	(Std. Deviation)	WFO	(Std. Deviation)
Syuzhet	0.6305	(1.2000)	1.5918	(1.2590)	0.9448	(1.2547)
Bing	0.5083	(1.4189)	1.6683	(1.3632)	1.4442	(1.4377)
AFINN	1.0717	(3.0767)	1.8892	(2.7325)	1.1733	(3.3849)
SentimentR	0.0730	(0.1933)	0.2012	(0.1749)	0.0970	(0.1552)

Table A2. Number of observations per work-arrangement subset

Work-Arrangement Dataset	Number of Total Tweets (Undersampled)	Number of Tweets from Average (Not Verified) Users	Difference
Work from Home (WFH)	1,200	1,140	-160
Work from Office (WFO)	1,200	1,099	-101
Hybrid Work (HYB)	1,200	1,095	-105

Table A3. Sentiment analysis of work arrangements of the average person

	WFH	(Std. Deviation)	HYB	(Std. Deviation)	WFO	(Std. Deviation)
Syuzhet	0.6341	(1.1998)	1.5891	(1.2590)	0.9379	(1.2547)
Bing	0.5237	(1.4189)	1.6712	(1.3632)	1.4385	(1.4377)
AFINN	1.1053	(3.0767)	1.8858	(2.7325)	1.1601	(3.3849)
SentimentR	0.0737	(0.8951)	0.2026	(0.1749)	0.0969	(0.1553)

Table A4 a. Impact of Elon's edict (EE) on WFH sentiment

WFH Sentiment	Before EE	After EE	(Std. Deviation)	Difference
Syuzhet	0.5931	0.5597	(1.2570)	0.0334
Bing	0.5021	0.5240	(1.4907)	-0.0219
AFINN	0.9997	0.7868	(3.0295)	0.2129
SentimentR	0.0699	0.1365	(0.1904)	-0.0667

b. Impact of Elon's edict (EE) on HYB work sentiment

HYB Sentiment	Before EE	After EE	(Std. Deviation)	Difference
Syuzhet	1.5882	1.3908	(1.1907)	0.1974
Bing	1.6569	1.5166	(1.3189)	0.1403
AFINN	1.8813	1.3676	(2.6518)	0.5137
SentimentR	0.2001	0.3745	(0.1745)	-0.1744

c. Impact of Elon's edict (EE) on WFO sentiment

WFO Sentiment	Before EE	After EE	(Std. Deviation)	Difference
Syuzhet	0.9314	0.7045	(1.2552)	0.2269
Bing	1.1425	1.1987	(1.4719)	-0.0562
AFINN	1.1483	0.5708	(3.1727)	0.5775
SentimentR	0.0957	0.1795	(0.1626)	-0.0838

Bryan Ji Yang Leong received his MS in data analytics from the State University of New York at Plattsburgh in 2022 and undergraduate degrees in politics and psychology from Messiah University in 2016. Since 2016, he has worked as a management consultant with Deloitte Consulting across the United States and Asia Pacific, where he currently is a senior consultant helping a variety of Fortune 500 companies solve their biggest problems. His consulting specializations include human capital strategy, technology transformation, organizational culture, leadership, and people analytics.

Zareef Mohammed is an assistant professor in the Department of Information Systems and Analytics at the McCoy College of Business at Texas State University. He previously held an assistant professorship and visiting lecturer positions at SUNY Plattsburgh and a lecturer position at the University of Miami in the Business Technology Department. He has received his MS in information technology and his PhD in information systems from Nova Southeastern University. Dr. Mohammed's primary research interests are in cybersecurity management and information privacy. He is also involved in merging neuroscience, in particular, electroencephalogram technology and theories, to better understand cybersecurity and information-privacy phenomena.