

# How K-12 School Districts Communicated During the COVID-19 Pandemic: A Study Using Facebook Data

Joshua M. Rosenberg  
University of Tennessee, Knoxville  
[jmrosenberg@utk.edu](mailto:jmrosenberg@utk.edu)

Ha Nguyen  
University of California Irvine  
[thicn@uci.edu](mailto:thicn@uci.edu)

**ABSTRACT:** The COVID-19 pandemic calls for urgent responses from school districts to allocate resources, develop instructional plans, and retain communication with students and parents. A channel to understand district responses is to examine how school districts communicate via public channels during the pandemic. Yet, understanding the nature of districts' communication when the pandemic was unfolding presents challenges, particularly when the analysis happens at scale. In this paper, we report the results of our use of public data mining, combined with text analysis techniques, to understand how United States school districts communicated about COVID-19. To do so, we compared district Facebook posts from 2019 to 2020 through topic modeling using 50,000 posts, randomly sampled from a corpus of 3,337,147 posts. We also determined how the topics we interpreted differed on the basis of a key school district factor, mean socioeconomic level. This work has implications for understanding K-12 educational institutions' communication about and during COVID-19 and the role of public data and learning analytics research methods for understanding these.

**Keywords:** public data mining, social media, educational institutions, macro big data

## 1 BACKGROUND

For communities across the world, the shutdown of schools due to the COVID-19 pandemic disrupted both students' learning and the tenor of daily life for students and their parents and caretakers. In this context, there were differences in how districts communicated with parents, students, and the wider community about their ongoing responses to COVID-19. Exploring this variation through digital—and public—data sources can provide an in-vivo lens (Salganik, 2019) into districts' communication as well as the myriad of ways the pandemic is impacting education.

To explore districts' responses, we took a *public data mining* (Kimmons & Veletsianos, 2018) approach, one that uses public data for research. Given the locally-controlled nature of education in the United States (U.S.), social media may be a channel through which districts communicate about many topics, from those related to enabling students to participate in schooling (e.g., posts about materials and devices) to celebrating students and others in the community. In this way, we employ an approach that is related to the use of *macro* (social media) data sources but it is different in that we focus not on one or a handful of institutions, but all that was accessible to us, and in how we use data from posts over many time points—akin to data collected in learning analytics systems (Fischer et al., 2020). Our research questions, then, are: How did U.S. school districts' communication via

social media change in response to the COVID-19 pandemic? And, to understand variation in their communication, how did this differ based upon an indicator of their mean socioeconomic level?

## 2 METHOD

Our data collection process, including the use of the CrowdTangle platform (CrowdTangle Team, 2020), which provides researchers access to Facebook data, is summarized in Figure 1.

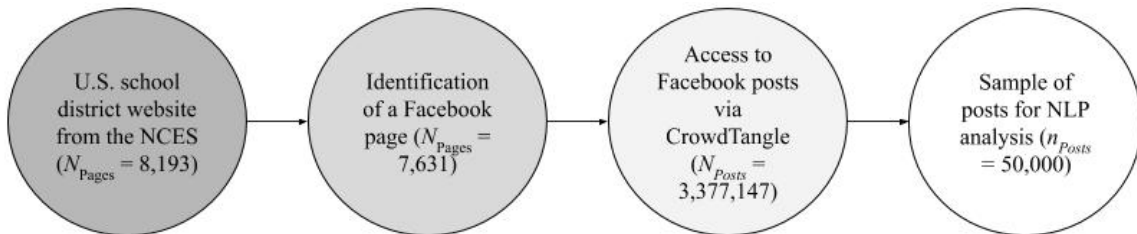


Figure 1: Data collection process used to identify posts from U.S.-based K-12 school districts

We then sampled 50,000 posts, due to the computational challenge of conducting NLP analyses on the entire corpus. The topic models, which were fit using the *stm* R package (Roberts et al., 2019), allowed topic prevalence to vary by year (2019 or 2020) and a measure of socioeconomic status (i.e., whether the percentage of students participating in a free or reduced-price lunch program, FRLP, is smaller than the sample’s average). To determine the number of topics *K*, we ran a series of models with *K* ranging from 10 to 40 and examined several model fit diagnostics, which suggested *K* = 20 indicate high predictive power and high semantic coherence (i.e., high-frequency words from a topic are likely to co-occur). The topic names (left panel, Figure 2) were determined by examining the top 15 high-frequency words and representative posts from each topic, followed by discussions between the authors.

## 3 FINDINGS

When ranked by frequency, topics about school spirit, help-giving, sports, registration, scheduling, and materials/device were those with the highest average prevalence in the corpus (see Figure 2).

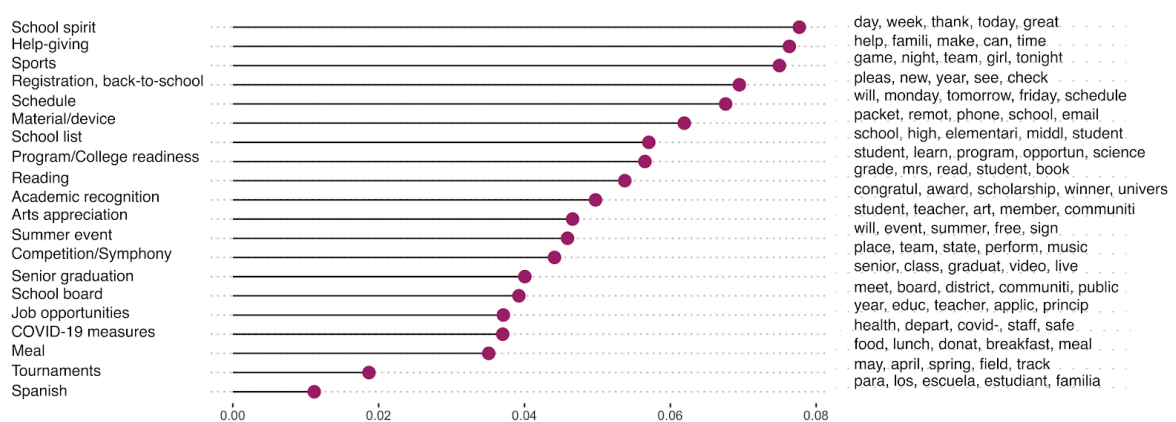
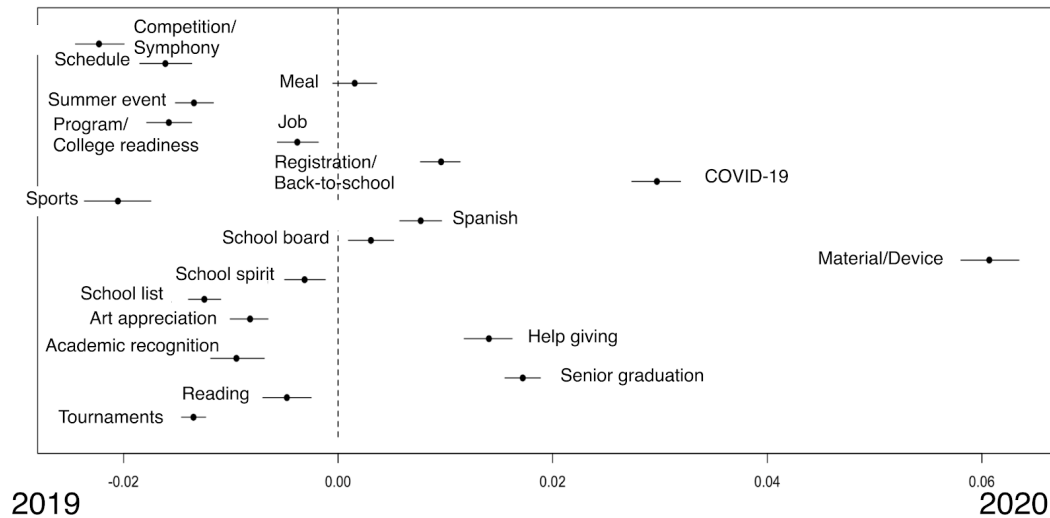


Figure 2: Topic Prevalence, Ranked by Highest-Lowest Proportion, and 5 High-Frequency Words

The topic models also allowed us to examine whether topic prevalence varied by covariates. Results suggest that in 2020, topics about registration/back-to-school, COVID-19, material/device, help

giving, and senior graduation were more prevalent (Figure 3). For example, on average, the proportion of topics about material/device pick-up was .06 higher in 2020 than that in 2019.



**Figure 3: Difference in Topic Prevalence between 2019 and 2020.**

Findings from the analysis of differences based upon FRLP suggest that overall, posts from school districts with a lower average socioeconomic level featured higher topic prevalence regarding meals, job posting, and materials/device pickup compared to those with greater socioeconomic resources.

## 4 DISCUSSION

This work demonstrates how a new data source—social media posts by institutions—can be informative about topics related to teaching and learning at scale. A key methodological feature that enabled this work was the identification of social media links on districts' websites. Indeed, we found that a surprisingly high proportion of districts may be using social media, but this data source has been examined very little in learning analytics research and educational research. Moreover, access to Facebook invites new learning analytics-driven questions (and ethical considerations, especially when posts by individual teachers or students are the focus of study, as they were in the present study), including those about variation in curricula and supports for students.

## REFERENCES

- CrowdTangle Team (2020). *CrowdTangle*. Facebook, Menlo Park, California, United States.
- Fischer, C., Pardos, Z. A., Baker, R. S., Williams, J. J., Smyth, P., Yu, R., ... & Warschauer, M. (2020). Mining big data in education: Affordances and challenges. *Review of Research in Education*, 44(1), 130-160.
- Kimmons, R., & Veletsianos, G. (2018). Public internet data mining methods in instructional design, educational technology, and online learning research. *TechTrends*, 62(5), 492-500. doi:10.1007/s11528-018-0307-4
- Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). Stm: An R package for structural topic models. *Journal of Statistical Software*, 91(1), 1-40.
- Salganik, M. J. (2019). *Bit by bit: Social research in the digital age*. Princeton University Press.