

Social Media as a Tool for Consumer Protection Monitoring

Dani Madrid-Morales
Assistant Professor
University of Houston

Melissa Tully
Associate Professor
University of Iowa

September 29, 2020



Outline

1. Why use social media data?
2. Context
3. Overview of the approach (methods)
4. Findings and examples from our project
5. Going forward



What can social media data provide?

- Monitoring of social media offers opportunity to collect **observational data** about consumers' opinions, attitudes and behaviors.
- Insights gathered from the **analysis of social media data** can
 - Help monitoring in real-time of issues/events
 - Be incorporated in policy interventions, A/B testing...
 - Be used in predictive modeling



Context

- We worked with Citibeats to conduct the social monitoring project
- Data collection driven by study goal:
 - Understand types of problems faced by **digital finance consumers**
 - Social media data collected in **Nigeria, Kenya** and **Uganda**
 - Data in multiple languages
- Data comes from Twitter, Facebook Public Pages, and Google Play Store
- Longitudinal study: from **July 1, 2019** to **July 1, 2020**
 - We cover roughly 6 months pre and post COVID-19



Approach/Methods

Step 1

- Collecting social media data at scale

Step 2

- Defining categories/topics of interest

Step 3

- Using word frequencies and probabilities to locate topics in data

Step 4

- Deeper analysis by incorporating user information



Approach/Methods

Collecting social media data at scale

- For this project, we collected **4.5 million social media messages** from
 - Commercial Banks
 - (e.g. Equity, Polaris Bank, Stanbic Bank...)
 - Telecommunication Companies offering mobile money services
 - (e.g. Telekom T-kash, Airtel Money, UTL...)
 - Fintech start-ups offering lending/payment
 - (e.g. Okolea, Sokoloan, Tala)
 - Microfinance institutions
 - (e.g. Uwezo Kash, Fortis Mobile Money, Tugende)



Approach/Methods

Collecting social media data at scale

Facebook
830,939 (42 %)

Twitter
1,651,659 (85 %)

. 23h
@KeEquityBank please assist. I can't access banking services through my eazzy banking app. I can't go past the login form after inputting my PIN.
1 Show more replies
Replying to @KeEquityBank
. 1h
I have done so
1
Equity Bank Kenya @KeEquityBank · 1h
We have responded, kindly check DM.

Equity Bank Kenya
1 d ·
Now grandma and brayo are sure that the only number equity will call them is 0763 000 000. when we call, we use that number only. Don't accept to be deceived.
#NumberYaEquity #OneEquityNumber Brayosammy
See original · Rate this translation
Equity Bank Kenya
158 78 comments 13 shares
Like Comment Share

Google Play Store
294,950 (40 %)

★★★★★ August 22, 2020 185
Branch!!! It's a good app, but my problem is that when u need money urgently, like for important issue, and u have the believe in mind that branch will give it out, that is when u will see, we can not approved your loan, check back in one month or two month time, and there's nothing to do about it, ...
Branch International August 22, 2020
umini, we are sorry to make you feel this way. Eligibility for loan approval depends on hundreds of factors that the system analyses each time you apply. For further clarification, please write to us through your Branch in-app chat by clicking on the "Customer Care" option on the app. VT

I am trying to instal your App and it keeps saying, "No connection." What could be the problem?
Like · Reply · 1d



Approach/Methods

Defining categories/topics of interest

- We combine **top-down** and **bottom-up** approaches to identifying salient categories in the data
- Through interviews and expert advice, **seven areas of interest** were identified before the analysis:
 - Operational failures, consumer care, fees & charges, fraud, data privacy, lending, advertising
- Using text analysis tools (cluster analysis), and human input (individual analysis of sample messages), sub-topics were identified



Approach/Methods

Using word frequencies and probabilities to identify categories/topics

- During the analysis, Citibeats used (semi-supervised) **machine-learning** to go from unstructured text data to structured numerical data

Initial seeds for category of **Fees & Charges** in Kenya:

- fees
- charges
- overcharged
- refund
- deduction

Step 1

User defined dictionary of keywords

17:28 - Jan 08, 2020

@ [redacted] Please return my funds to my account. Yesterday i had a balance today it negative. Please what happened. Please refund my money.

Step 2

Computing topic probabilities from keywords & context

18:59 - Jul 30, 2020

@ [redacted] Having Nyeri1 return my money should be as simple as it was for them to craft the false statement.

Step 3

Newly learned words help determine topics for items with no keywords

14:30 - Jul 30, 2020

@ [redacted] @ [redacted] This bank if you don't follow up hiyo pesa itaogelea



Approach/Methods

Deeper analysis by incorporating user information from metadata



A screenshot of a tweet from the Twitter API (@TwitterAPI) dated October 10, 2018. The tweet text reads: "To make room for more expression, we will now count all emojis as equal—including those with gender and skin tone modifiers 👍👍👍. This is now reflected in Twitter-Text, our Open Source library." Below the text, it says "Using Twitter-Text? See the forum post for detail:" followed by a link to a forum post titled "New update to the Twitter-Text library: Emoji charac...". The tweet has 297 likes and 188 retweets.

```
"place":
{
  "attributes": {},
  "bounding_box":
  {
    "coordinates":
    [
      [
        [-77.119759, 38.791645],
        [-76.909393, 38.791645],
        [-76.909393, 38.995548],
        [-77.119759, 38.995548]
      ]
    ],
    "type": "Polygon"
  },
  "country": "United States",
  "country_code": "US",
  "full_name": "Washington, DC",
  "id": "01fbe706f872cb32",
  "name": "Washington",
  "place_type": "city",
  "url": "http://api.twitter.com/1/geo/id/0172cb32.json"
}
```

Location

User device

Gender

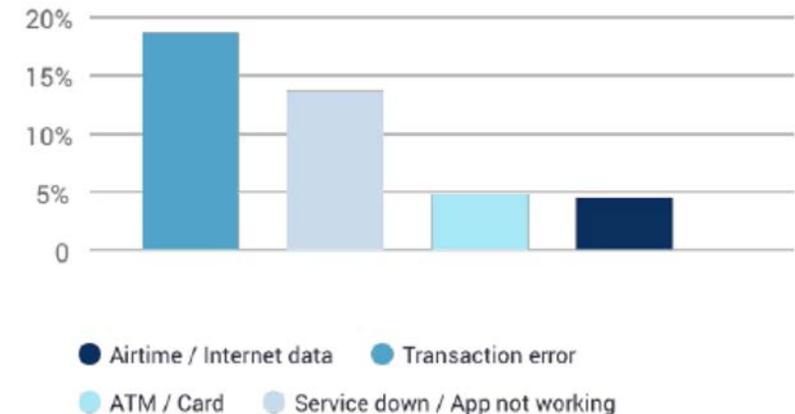


Preliminary Results

These results are subject to change.

1. Twitter and Facebook are mainly used to report **consumer protection issues, particularly customer care**
2. Google Play Store reviews focus on **app performance (positive reviews) and operational failures (negative reviews)**
Transaction errors are the most common operational failure reported

Types of operational failures reported on Google Play Store
Proportion of comments related to different types of issues.
Three markets aggregated



Preliminary Results

Response rate proxy on Twitter, Facebook and Google Play
Number of responses of banks / Total of tweets, by type of issue

3. Financial providers' response rates **vary considerably** across Twitter, Facebook and Google Play Store

4. Replies on Twitter are more **concentrated on customer care** issues; Facebook and Google Play responses are **more distributed among different issues**



Preliminary Results

Results related to Covid-19

5. The use of social media channels to communicate issues and interact with financial providers **increased** across the three markets after the Covid-19 pandemic (early March 2020)
6. The distribution of issues **did not change** post-Covid-19
7. However, on Twitter, women in Nigeria have **significantly increased** their rate of complaints about fraud compared to men, while in Uganda customer care reports have also **risen** for women



Accessing and Analyzing Social Media Data

My Datasets

My Projects



Categories

13/02/2020 - 31/03/2020

Datasets & Data sources

search by term or expression (eg: term1 + term2)

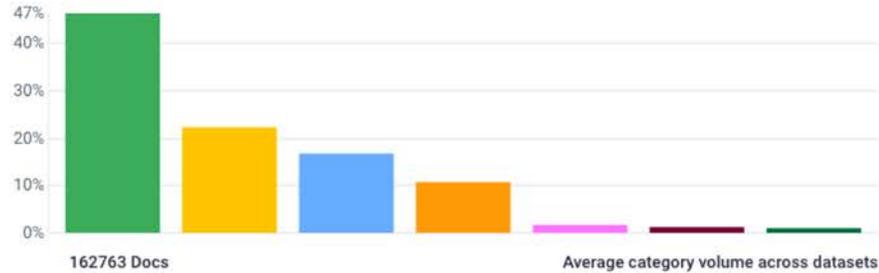
Select all

- 01 Operational Failures
- 02 Customer Care
- 03 Fees & Charges
- 04 Fraud
- 05 Data Privacy
- 06 Lending
- 07 Advertisement

Volume across datasets

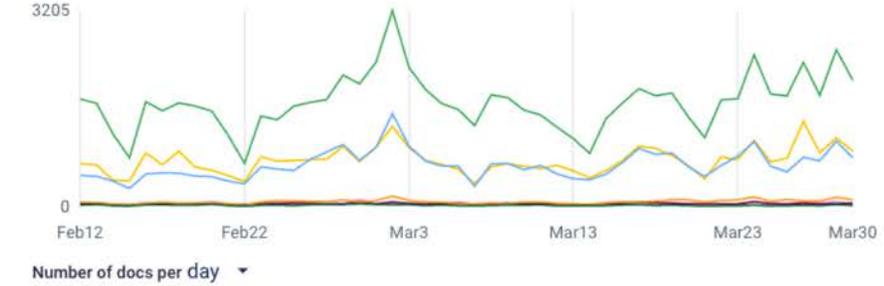
What is the priority overall?

VOLUME SENTIMENT



How was the evolution?

ABSOLUTE RELATIVE



Datasets Overview

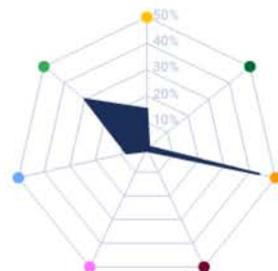
Click on a dataset to explore it further.

KE_Commercial Ba... explore



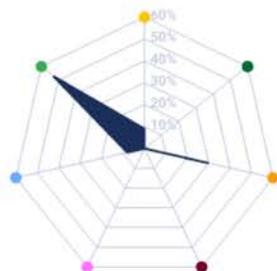
12971 Docs

KE_Fintech_T explore



360 Docs

KE_Microfinance_T explore



47 Docs

KE_Telecomms_T explore



39612 Docs



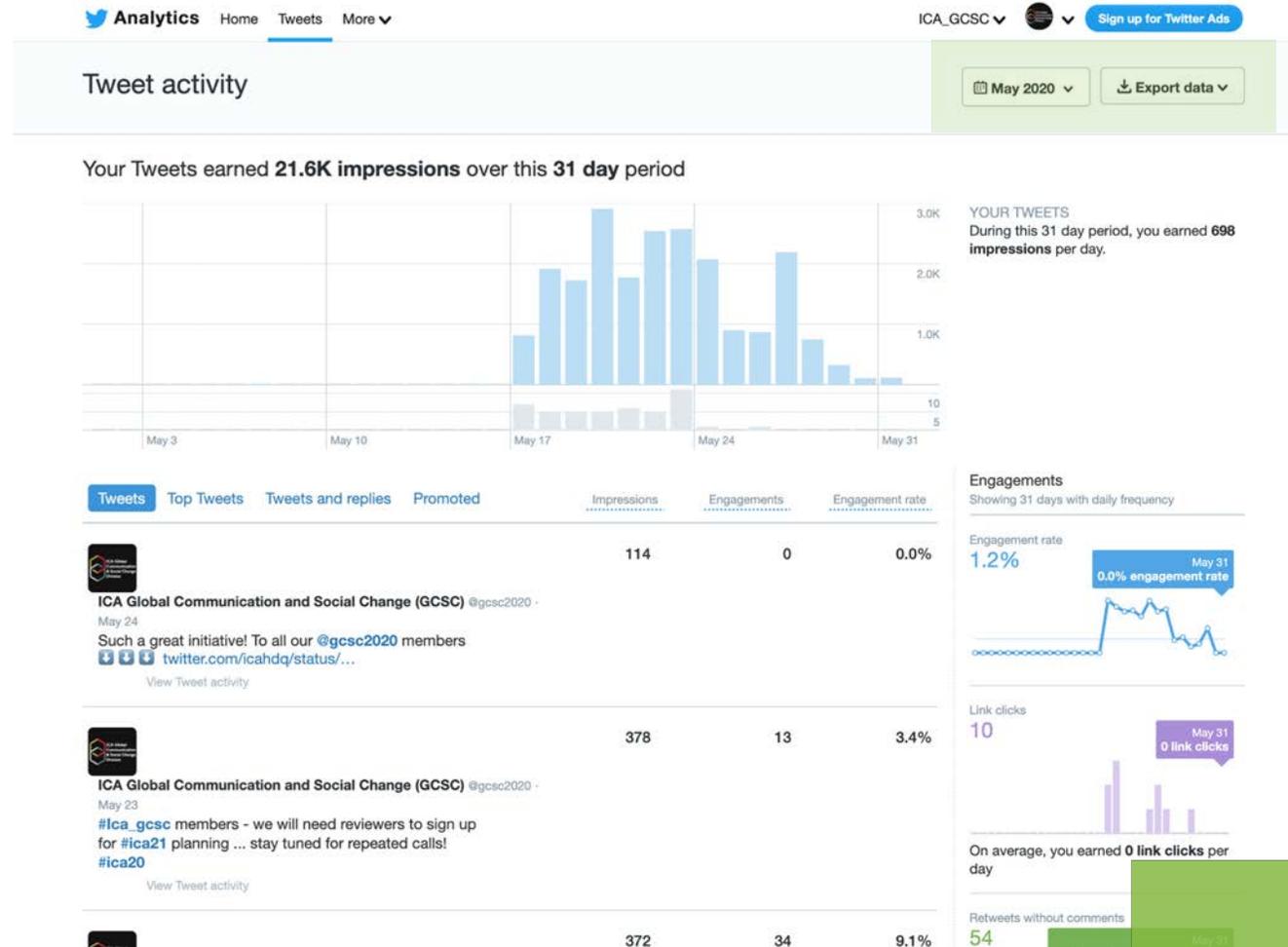
Tools for Accessing and Analyzing Social Media Data

- SM platforms have analytics dashboards
- Lots of external tools/services available but...
 - Some are very expensive
 - All have limitations



Tools for Accessing and Analyzing Social Media Data

Free tools

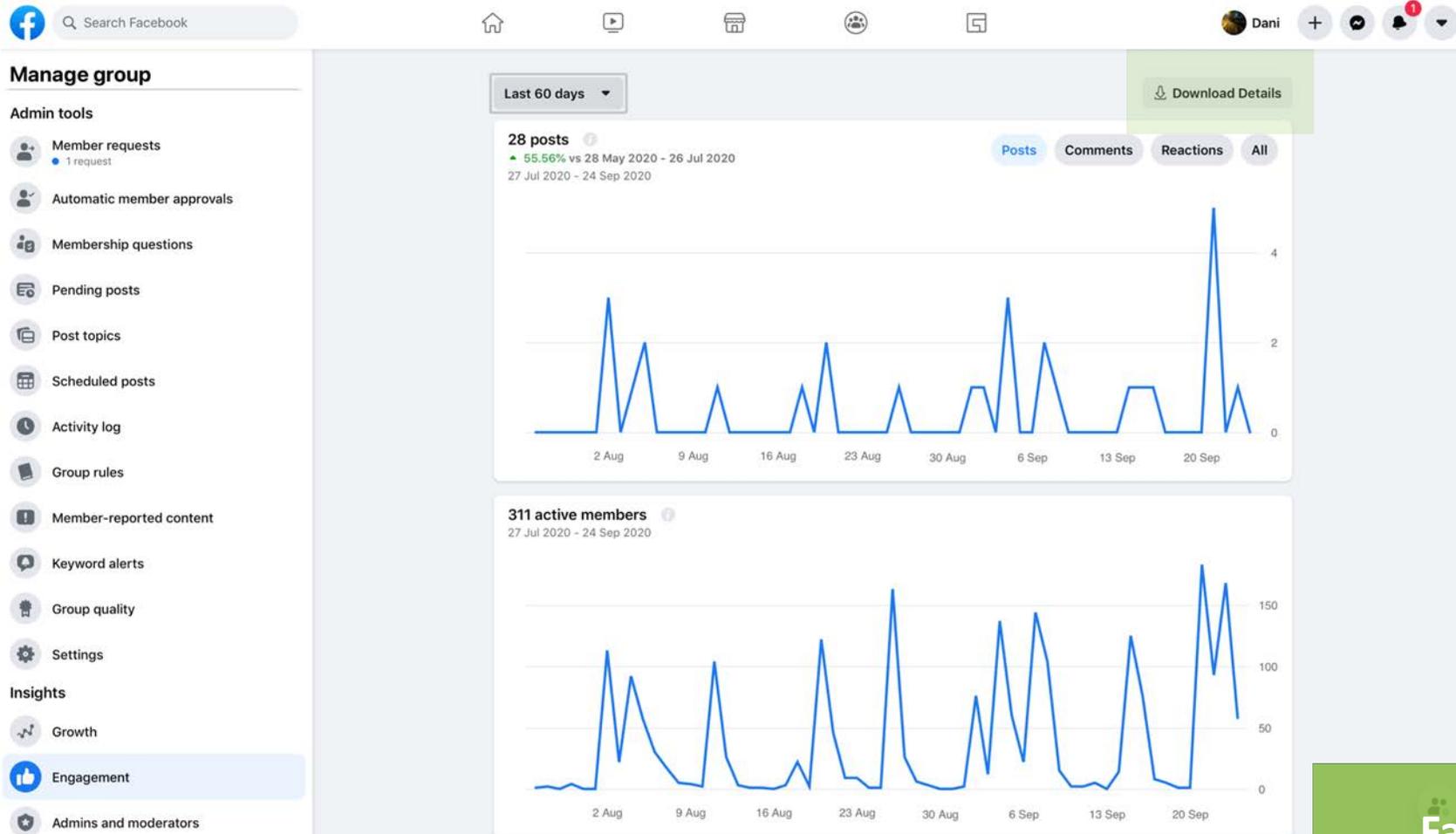


Twitter Analytics



Tools for Accessing and Analyzing Social Media Data

Free tools



Facebook Engagement



Takeaways

- On a large scale, social media channels offer a **rich source of consumer protection data** that can be gathered and analyzed to learn about customers
- On a smaller scale (e.g. at the individual bank, provider, regulator level), **data can be collected** and analyzed to glean relevant insights
- Data reveal patterns in both complaints and responses that can be used to **inform decision-making**



Thank you



ipa
INNOVATIONS FOR
POVERTY ACTION

Special thanks to Citibeats for their work on gathering and analyzing the data.