

Social Media as a Tool for Consumer Protection Monitoring

Q&A

The questions below were received electronically during the First Practitioner's Forum Meeting of IPA's Consumer Protection Research Initiative on Sept. 29, 2020. These responses were prepared by Laura Kirchner, Alicia Leon, Dani Madrid-Morales, Rafe Mazer, Melissa Tully, and Harry Wilson. The comments below reflect the responding team's views and not those of Innovations for Poverty Action.

A. Greenberg: Good morning. I have a question. For the application for funding to test solutions to consumer protection problems in the digital space, does the applicant have to be based in one of the four countries prioritized by the project, or could an FSP from a different country from Bangladesh, Kenya, Nigeria, or Uganda also apply?

The current priority countries are those four, but if there are particularly relevant opportunities outside those countries there may be room to consider. I would recommend reaching out to our team to discuss the idea informally for feedback. You may write to us at financialinclusion@poverty-action.org.

B. Magnoni: I have a question. How do you ensure that the consumer feels "heard" when you are working with the data on the back end and not responding and interacting in real-time?

We have seen some financial service providers using chatbots (these can be relatively easily set up on Facebook) to automate responses. So, if a provider wants to monitor user satisfaction/complaints automatically, they could provide automated responses to reassure the customer that they are being heard, while the issue is escalated to whoever can address the concern. Also, we have seen a few financial providers reuse set responses that are probably "copied and pasted" onto social media every time a consumer expresses some concern or complains about an issue.

M. Griffin: Did you pick up any non-individual tweets, e.g., consumer associations, gender groups

These groups (consumer associations, gender groups...) would be included in our dataset. The way we collected the data is to capture all mentions of digital finance service providers. So, if we are thinking of Twitter, that would be all messages (from individuals, organizations, groups...) that would have addressed an institution using the @username. On Facebook and Google Play Store, we collected all messages directed to the financial institutions. While we collected these data, we did not analyze the interactions between these groups and financial institutions specifically.

L. Graham: Are all the posts tagged to one of the financial institutions, or does this include people discussing similar matters amongst themselves?

On Twitter, to retrieve the data, we identified relevant posts by looking at tags, so if two users discuss services by provider A, but do not tag that provider, we did not include that post in our dataset. The reason for not including these, is that using tags creates a systematic way of collecting the data. If we were to look for instances of providers' names, we might find that some are misspelled or abbreviated, and it would be difficult to have consistency in the data. There also may be cases where the name of the provider has many unrelated associations. For example in a 2018 analysis of Twitter data in Kenya, the name of CBA, one of the larger banks in the country, also stood for "Can't be asked" and so just monitoring for use of "CBA" without the @ created lots of irrelevant content. When using large volumes of data there's always a trade-off between comprehensive and consistency. We know that we might have left out some messages, but given the size of our dataset, we can approximate general patterns rather accurately.

I. Aprilianti: Hi, excellent method and findings! I'm Ira from Indonesia. I have a question. Of course, in the view of consumer protection, we should ensure not to use personal protected/private data. However, in the view of research, wouldn't it be selection bias if you only include public posts? If it is a selection bias, how would you minimize it?

For ethical reasons, we are only able to include public posts, but we acknowledge that this limits the data and does not include all possible exchanges between consumers and providers. We also acknowledge that social media, by its very nature, is limited to only those consumers who have access to the platforms, which also skews the data.

T. Kumari: In Asia, even LinkedIn is being used also for complaining and sharing feedback on Financial Institutions too.

We haven't collected data from LinkedIn for this project, but this is definitely a possibility of a type of social media data that could be added to a similar analysis. Each country might have different prominent social media. In the case of the countries we are studying, we did not see LinkedIn being so prevalent, but it is worth considering.

S. Zehra: Did your analysis pick-up whether social media was the first channel used by the users or was it a second channel or an alternative channel? Basically, was social media their first choice to make complaints?

We did not include this question in our analysis, but this is definitely an interesting area of research that we might be able to explore in future analyses of the dataset.

M. Griffin: FSP responses - can you tell whether they are auto-generated or manually, staff generated?

Although we cannot be sure if the responses are auto-generated or manually generated, we can look for duplicates in the responses to see if the same phrases are used repeatedly. This would indicate that either an auto-reply is set to respond to certain keywords/inquiries or that the institution has a standard set of responses that staff provide.

M. Mishra: Do you have labels which pick the sentiment or gradient of the Responses (0/1) or a simple ordinal classification (Eg: 0-5)

Our classification of sentiment is categorical, meaning that we have closed categories (positive, negative, neutral), not intensity levels.

T. Kumari: But twitter may be less preferred channel for microfinance because of better access and response from using MFI Apps which is part of Google Play Store?

This is a significant limitation for monitoring consumer complaints to FinTechs, as many customers will use the in-app customer service tools. This is where complementing social media data with customer care logs from providers would help to complete this picture. That is exactly what we will be presenting in our next webinar in late October, and both research teams are in fact discussing how to use the social media and customer care data to complement each other.

P. Nelson: Did you get insight at all on gender dimensions to these complaints raised or problems reported?

During the presentation, we only presented gender differences in pre- and post-COVID 19 consumer experience, but in the final report that will be made available once the research is completed, we were able to examine other differences between genders. All our gender-related analysis is based on Twitter data only. The data collected from Facebook and Google Play Store does not include metadata about users.

D. Dias: I see a lot of value in monitoring social media for many purposes in CP supervision by financial authorities. But I am wondering why a financial sector supervisor would use social media monitoring specifically for complaints, when they can get access (real time if they want) to all channels used by FSPs to address complaints, including their social media channels. It only doesn't work if the FSP does not register these channels in their

complaints database. But that is another issue (bad practice that could or not be linked to the lack of regulation on this issue).

This is a very interesting point. It might be possible that some users are not familiar with the processes to complain directly to FSP and, therefore, do not use traditional channels. In other instances, the monitoring of social media data might be a way to foresee problems that haven't been escalated to customer care representatives (and therefore not included in complaints databases), because users did not think they were "too serious", or because providers might not see the need to report those. Another situation that social media would capture is when a consumer is not satisfied with a complaint outcome so goes to social media. There are postings noting a case that is severe and has gone a year without resolution. Those cases may be logged as pending with no further details or even incorrectly deemed "closed" by the provider. Regarding the point about accessing the channels used by FSPs in real time, I believe that is still not common practice in most markets, so this offers an intermediate step between that and the current reporting of complaints data by providers on a monthly or less frequent basis.

In parallel to this study, IPA is working on another study that uses complaints data from FSP submitted to regulators in Uganda. One of the follow up projects that we are considering working on is to compare the types of complaints reported in the databases, and the types of complaints Ugandan users express on social media. This might help us better understand the point raised in this question.

J. Chien: Can you link the data to specific products? Analyzing trends in issues by product and provider would be very useful.

Yes, we have the data for specific products and services so we can see if certain issues arise related to specific products/services offered by an institution.

P. Okull: What are the IRB requirements for using such data?

All the data used for this study is publicly available data. It is the equivalent of letters to the editor in a newspaper, or comments posted on the website of a newspaper. When users post these messages, they are aware that they will become publicly available. Because we did not collect any additional user data other than what is publicly available, no special IRB requirements are needed.

A. Greenberg: Did you do an analysis of the types of phrases or manners of expressing dissatisfaction that seemed more likely to get a response from a provider?

We did not go into that level of analysis in this report. However, this is an interesting question that we could pick up in later stages of the project. One possible way to

approach this would be to compare responses based on the "sentiment" of users (positive, negative, neutral), and test whether negative sentiment gets more responses. We will definitely take this suggestion and see if we can explore it further.

A. Namubiru: Is there a way to incorporate the DM tweets as well in order to balance the 55% DMs conversations reported so as to have a more balanced % on the twitter finding??

We cannot include the DMs because they are not publicly available. The only way to do that would be to get the messages directly from the institution or individual who was involved in the exchange.

D. Dias: Did you do, or could you have done prioritization of the consumer risk based on this analysis? That would be useful for supervisors.

We think so. The different categories of problem types are the main aspect we would organize this prioritization around. But what is nice about this tool is you could then look at each of those issues filtered by time--to see spikes during certain days or weeks--by products and by providers, so you could then follow-up on the most common issues not just at market level but on a provider by provider basis. In a prior study with CitiBeats, we tested an alert system where any deviation of more than 10% of the rolling monthly average of social media comments on consumer protection to a provider triggered an email with a link to that data set to investigate further.

M. Griffin: Is there any way to track by source to determine persistency of issue, e.g., response doesn't address concern?

Are you referring to a consumer raising their issue several times due to non-response? That can be done by analyzing by consumer profile to see how many consumers have tweeted multiple times. This would also help flag influencers who amplify individuals' problems by reposting them.

T. Kumari: Are all the social media complaints or feedback integrated with the central complaint tracking system of MFI as well?

We do not know if MFIs have integrated these into their complaint tracking system. In visits to customer care centers of larger entities like MNOs and banks we have seen specific social media desks in the call centers who handle complaints via those channels.

P. Nelson: I'm also curious to what degree this type of data has led supervisors to issue circulars or alerts. As opposed to issuing such circulars based on data already collected from supervised entities.

We are hopeful that the data will be used in this way - to inform how institutions respond to consumer protection issues on social media and beyond. We believe that social media data offer insights that can be used to develop messages/warnings and procedures to address these concerns.

P. Nelson: (My question gets to my curiosity of how much credence that supervisors now place on social media-sourced data. If it's not considered (correctly or not) high-quality or representative, then we might not yet supervisors relying on it in this way.)

Anecdotally, one supervisor we have spoken to about this work noted the bias and that they were not concerned about this, as they find that more educated or proactive consumers may flag issues that are pervasive with other consumers as well, so they can be a "canary in a coal mine" to an extent. There is a question about enforceability of course. You would not punish or mediate a dispute resolution off of a Facebook post. There would certainly be need for more investigation, but that would also be true of issues flagged in customer care logs or complaints presented directly to the regulator. One unexplored aspect of this channel is you can easily send a private message a consumer to see if their issue was resolved and gather further information without having to speak on the phone or receive them in person.

F. Gwer: Do you have a breakdown of response rate per keyword?

At this point, we do not have the breakdown by keyword. Given the way we have classified our data, it would definitely be possible to see whether positive, negative or neutral messages get higher or lower response rates. In addition, it would be possible to compare different topics (e.g. customer care vs. fees and charges).

Also, in a sense, this would be similar to using keywords given the way we have come up with the different topics. For each topic, we have created a list of "seed words" (or keywords) and used these keywords to identify what topic each post is about. In addition, we could compare posts by which keyword is featured (not by topic), but this might not provide as much insight.

We will try to incorporate these ideas in the first round of analysis and might be able to present some of these findings when the initial report comes out.