

## Further Analysis and Possibilities with Respect to Anonymity in Social Media

Erik Sandewall

The use of anonymity in social media has been studied and debated during the last ten years [1] , but this question comes up with renewed vigor from time to time, most recently in the U.S. where presidential candidate Nikki Haley has taken a strong position on it [2] . An early article from the Pew Research Center contains a comprehensive analysis of the issues and the opinions that apply to this topic [3]

Most of the articles in [1] take a black-or-white approach to this question whereby some adopt the view that anonymous postings should not be allowed at all in social media, and the others (the majority) propose that all authors must be allowed to be anonymous if they wish so. However, when their accounts are taken together, one must conclude that both may be right, in the sense that there are some situations where it makes sense to allow anonymity, and some other where it does not, provided that reasonable editorial policies are applied in both of these cases. The purpose of the present paper is to discuss how such policies may be defined.

### Some On-line Environments are 'Social', others are More like 'Media'

The term 'Social Media' refers to on-line systems whose users may submit statements of their own opinions or knowledge, read the statements of other users, and add comments about the statements (or comments) of other users. In the present context, I propose to make a distinction between those social media that harbor discussions, and those that harbor an on-line counterpart of mass meetings. In one variety of *mass meeting media* there is one or a few *leaders* whose posts are read by a very large number of *followers*, and where many of these followers are active by contributing comments to the leader's messages, usually comments of agreement and support.

In a *discussion media* there is no such distinction between leaders and followers, and the purpose of the interactions is to exchange information and opinions between the participants. One of the participants may be designated as the 'chair' of the discussion, but otherwise the participants are more or less on the same level.

---

Notes:

---

Usually, mass meeting media will be open for anyone who wants to attend, without registration or with a facility for easy on-line registration. On the other hand, discussion media will typically be closed-shop events where the members join by invitation, or after an application procedure. Moreover, what is said during the session of a discussion media may be assumed to stay with the participants, in many cases, whereas mass-meeting media are usually intended to reach as many people as possible, even after the session has ended.

On-line classrooms with a one or a few teachers and a larger number of students will count as discussion media in this sense, even though there are different roles for the teacher and the students there.

We shall first use the simple distinction between mass-meeting media and discussion media and discuss the pros and cons of anonymity in these two cases. An additional case that does not fit readily into the initial distinction will be discussed towards the end of the article.

### **Anonymity in Discussion Media**

Participation in a discussion media can usually be assumed to be based on each person's explicit wish to participate. However, there are also some situations where individuals are more or less obliged to participate in such a group, for example when a school organizes a discussion group for the students in a particular class. I shall refer to these as the *voluntary* case and the *assigned* case, respectively. In both cases I shall also assume that attendance is limited, and that the contents of the discussion are not supposed to be released to the public.

In the voluntary case, which must be the most common one by far, it would seem natural to use proper names in such discussion groups, but some cases have been mentioned where pseudonyms would be preferred. For example, in some teaching situations one may organize a discussion where one-half of the participants are explicitly asked to adopt one political position for the sake of the exercise, while the other half are asked to adopt an opposite position. The statements that are made by a participant during such a discussion must then be distinguished from their real opinion, which would be a reason for using acronyms.

The conclusion would be, therefore, to leave the anonymity issue to the discretion of the organizer of the discussion media. If someone does not like it then they are free to leave.

The case of assigned participation is a more difficult one. It is illustrated by a case in Irish court [4] involving the Salesian Secondary College, a school in Limerick whose board of management is in dispute with Facebook Ireland. The school wants Facebook to disclose the identity of individuals behind an Instagram account in order to enforce its own social media code of conduct. Facebook says it can not do so without a court order or a request from law enforcement.

---

In such situations there may be reasons to breach the anonymity if bullying and harassment are suspected to occur in the discussion, but there may also be reasons to protect the anonymity, for example if teachers have been criticized in that discussion. One possible solution may be to nominate an external person who is allowed to see the current contents of the discussion, and who would have the power to deregister participants that have behaved inappropriately. This reviewer should not be associated with the staff of the school, of course.

Similar situations may possibly occur if a company has set up an assigned participation group for its employees. In this case it may be natural to ask the *skyddsombud* (an ombudsman for the employees) to do the review.

### **Anonymity in Mass-Meeting Media**

One variety of mass-meeting media was described above, as a situation where one or a few leaders are able to broadcast their messages to a large number of followers, and where the latter are invited to add their comments, but where disagreement and debate is not really expected. However, another variety may occur if each session consists of a large number of *threads*, each of which is initiated by one participant and followed by comments by other participants. Both these varieties may occur irrespective of the choice of platform; both on Facebook and on Twitter/X for example.

Many arguments in favor of banning anonymous statements in social media see this as a means of dealing with two problems: the occurrence of hate speech and harassment under the protection of anonymity, and the manipulation of public opinion through the use of automatic message generation or so-called trolling. Neither of those problems is likely to be of significant importance in the multi-thread variety of mass-meeting media, with one exception that will be discussed below. It is appropriate, therefore, to focus on the 'rumour' situation where a false statement is forwarded to a large number of users by being quoted and requoted in a process of exponential growth. It would seem that the leader/follower situation is particularly favorable for that kind of growth, in particular if it means that followers have a habit of returning frequently for reading recent news that are related to their leaders.

There is no single solution to the problem of the 'rumour' situation, as we shall see. However, the use of user validation whereby each user is known to be an actual person with reasonable properties will at least go a long way towards eliminating the use of 'bots'. This applies regardless of whether users are identified by name or by pseudonym.

The use of proper names does not in itself help with the 'rumour' problem. One recent example from Sweden may illustrate how this can happen. During an informal meeting in Gothenburg on November 21, prime minister Ulf Kristersson

---

was interrupted by protests when he was speaking about the Israel-Palestina conflict. No violence ensued, but eight of the protesters were requested to leave the meeting. Shortly after, there was a post on X that identified one of these eight by name, stating that he was a known member of a criminal gang. This was incorrect – it was a clear case of fake news – and the person in question was a 22-year student who was born in Palestine and who is a vice-chair of the European Palestinian Youth Union, and without any criminal history whatsoever.

The author of the post on X withdrew the tweet when he was notified of the mistake, but by then it had already been seen by many. Worse still, one of the major national newspapers – *Svenska Dagbladet* – picked up this information from the tweet and included it in one of their editorials, without having checked it sufficiently. It is estimated that the fake news may have reached several hundred thousand persons through the combined effects of X and the newspaper story. This illustrates the speed and the ways by which a piece of fake news can spread. In this particular case the error could be corrected relatively quickly, although still causing considerable discomfort for the students involved.

There are many other situations where the error is not so easily distinguished and where the surrounding arguments may confuse the issue, which makes it even more difficult to check session contents for correctness. This applies in both of the varieties described. Furthermore, there is a problem of scale: the number of contributed comments will often be so large that effective supervision of their contents is not an option, realistically, at least if the vetting must be done by human operators. Artificial Intelligence techniques have been proposed and tested for this purpose, but they may be easy to circumvent. It would be ideal if incorrect postings could be caught before they are allowed to spread, but this is not realistic, therefore.

Several other approaches for solving the problem of incomplete supervision require that contributions are published with the proper name of the author, applying both to the comments and to the thread-starters. Applying the force of the law, by taking the authors of reproachable contributions to court is a theoretical possibility, but it is unlikely to succeed except in very exceptional cases.

Another approach may be to rely on shaming, either immediately after the time of publication, or in an expected future. Immediate shaming would mean that references to reproachable contributions are added to a separate blacklist, with a similar blacklist that just shows the authors that have made blacklisted contributions. Unfortunately, one may fear that some participants would view placement on such a blacklist as desirable in itself.

Future shaming, on the other hand, would rely on an assumption that participants would be wary of posting statements that could have a negative impact on a prospective future employer, or even a future spouse or father-in-law. In other words, the possibility of future shaming might have a restrictive effect on what

---

people choose to write today. However, it is not at all clear whether this factor would have much effect.

One particular problem with shaming in general, and future shaming in particular, is that there may be more than one individual that has the same combination of first name and last name, even including initials or other components of the full name. One solution to this problem was proposed in a previous article [5] .

In spite of its limitations, shaming mechanisms may help to some extent for reducing both the problem with hate speech and harrassment, and the problem with the manipulations of opinion. This may be used as an argument in favor of using proper names rather than acronyms. Moreover, our sense of justice dictates that everyone should take responsibility for their actions, and this ought to be a moral argument for the use of proper names.

### **Counteracting the Actual Spread of Fake News and False Statements**

In an ideal world, both hate speech, fake news, and other kinds of disinformation should be captured by filters before they get to be published, but such a goal is not achievable. As a second-best, one may consider methods for controlling the spread of such disinformation once it has started to happen.

All such measures must be based on the dissemination of reliable facts, and on the dissemination of alternative viewpoints in those cases where there are legitimate disagreements. One far-reaching measure in this direction might be to establish one or a few 'authorities', in the sense of organizations that promote reliable facts by collecting, organizing, and publishing them in readily available form, at least for facts in some selected areas of particular importance. If such authorities could be created then they would be useful in all situations where incorrect information is spreading, either due to ignorance or due to malevolence.

In particular, one could foresee an automatic system that monitors messages in mass-meeting media and that are able to identify posts that seem to contain disinformation. Such a system should then react by sending a message to both the originator and the recipient(s) of the post in question, encouraging them to look into the authority coverage of the topic in question. The message should never be saying "you are wrong"; it should merely point at relevant information and encourage the users in question to look into it.

Authorities of the kind described here could be set up as new and independent organizations, but they might also be organized within universities of high standing. Performing such a service would arguably be a very natural part of the universities' mission in society.

---

## **Counteracting the Spread of Disinformation by Ticketing**

The arrangements that were described in the previous section may be useful for counteracting disinformation that spreads at moderate speeds, but they may be much less effective against rumours that spread very rapidly after having been originated, as was the case after Kristersson's meeting with the public. One other possibility remains, namely the post-fact discouragement of the spreading of rumours, by administrative means.

The current practice of assigning 'tickets' for parking violations may be seen as a precedent for this approach. The fact you were in good faith when you parked in a particular place is not a viable reason for refusing to pay the ticket, and the same holds if you were not able to return to your car within the allocated time, for reasons beyond your control. Parking violations are so frequent, and their consequence is so moderate that it would be impractical to use ordinary court procedures in every such case.

In the same way, one might institute a system whereby a 'fake-news ticket' could be issued to everyone that has forwarded a piece of information that turned out to be fake, in particular if this has been done using social media whereby the action can be documented fully. Actually, the same approach could be considered when the fake news is forwarded (or 'shared') using a messaging system, including electronic mail.

The overriding purpose of such an arrangement would be that people should think twice before 'sharing' a piece of news, and they should check their source carefully before doing so. There would be some collateral damage, of course, since people will sometimes forward a piece of fake news although they are in good faith and believe that their source is absolutely trustworthy. Getting a fake-news ticket would simply be one more nuisance in life for those who like to share information with their friends.

## **Conclusion with Respect to Anonymity in Mass-Meeting Media**

From all these considerations, I would argue that the use of proper names is to be preferred over the use of pseudonyms in Mass-Meeting Media, especially from the point of view of society. The use of proper names does not solve all problems, but it does facilitate several ways of mitigating them.

Accordingly, it seems natural that each operator of such a media should be able to choose their own policy with respect to anonymity, as a natural part of their general editorial policy. If a media has decided on a proper-name policy, then a user should not be able to demand anonymity, even under a free-speech argument.

---

## Opportunities for Making Controversial Proposals Anonymously

One of the situations that requires particular attention is when a person wishes to make an original and radical proposal with respect to a question that is important for society, when this person is also concerned about their own security since the proposal may be so controversial. There is a traditional solution to this problem, namely, to submit the proposal for publication in an established newspaper or other periodical publication. This approach will at least work in countries that have adopted the system of 'responsible publisher' as described in [5] . Under that system, the responsible publisher is the only one that has to appear if some of the journal's contents are challenged in court, and the publisher has a legal right to protect their sources, and an obligation not to divulge them. This is an established system for paper-based publications, and it can easily be extended so as to apply for electronic periodicals as well. In fact, the present article is published in the electronic periodical *Liberal Kommentar* which has been registered in this way under Swedish law.

Some well-known and much earlier cases of this kind have occurred in the United States, as described by [6] :

*Anonymous speech is a hallowed American tradition. It's hardly an exaggeration to say that if it had never been permitted, the United States as we know it may never have existed. Letters from a Farmer in Pennsylvania, secretly penned by John Dickinson, were instrumental in uniting colonists against British misrule. Alexander Hamilton, John Jay and James Madison published their stirring Federalist essays defending the proposed U.S. Constitution under the pseudonym "Publius." Thomas Paine, Benjamin Franklin, Thomas Jefferson regularly penned missives without their real names attached.*

These events are hardly relevant for us today since the use of a newspaper was not an option in the eighteenth century; they hardly existed at all at the time, and they were certainly not in general use. This distinction is further confirmed when [6] continues:

*For good reason, the tradition has remained alive through the 21st century. In 2018, the New York Times published an op-ed from a then-anonymous official who proclaimed that "the president continues to act in a manner that is detrimental to the health of our republic." Donald Trump, of course, said it was treason.*

The uncontrolled use of anonymous postings in social media in the name of free speech can hardly be defended by examples such as these. The 'anonymous official' did choose to publish their op-ed in a major journal.

---

## Controversial Proposals in the Present Article

In the course of discussing possible measures for counteracting perceived negative effects of anonymity, the present article has mentioned several ideas that are likely to meet with strong objections. This applies, in particular, for the idea of 'authorities' that could serve as the guardians of true facts, and the idea of issuing 'tickets' for having shared fake news. It is important to emphasize that these ideas are being mentioned here in a spirit of brainstorming; they may stimulate a discussion of their pros and cons, and they may also serve as inspiration for coming up with other unconventional approaches to the very real problem of disinformation in our society.

## References

- [1] Annex: *List of Articles on Anonymity in Social Media* .  
<https://www.liberalkommentar.se/articles/257/annex.html>
- [2] *Nikki Haley walks back her demand that social media ban anonymous posters after facing GOP backlash*. Associated Press, 2023-11-15.  
<https://apnews.com/article/nikki-haley-anonymous-social-media-online-8fa8bfd5d9022f3fa2d06b8f474587d6>
- [3] Lee Rainie, Janna Anderson and Jonathan Albright: *The Future of Free Speech, Trolls, Anonymity and Fake News Online*. Pew Research Center, 2017-03-29.  
<https://www.pewresearch.org/internet/2017/03/29/the-future-of-free-speech-trolls-anonymity-and-fake-news-online/>
- [4] Ann Henry: *EU court asked to rule on right to online anonymity*. Pincen Masons, 2021-05-26.  
<https://www.pinsentmasons.com/out-law/news/eu-court-right-to-online-anonymity>
- [5] Erik Sandewall: *Does Freedom of Expression Entail a Right to Anonymous Publication?* Liberal Kommentar, 2023-11-27.  
<https://www.liberalkommentar.se/artiklar/256/freedom-and-anonymity.pdf>
- [6] Editorial Board: *Nikki Haley, Big Sister: Speech, even anonymous speech, must be protected, not banned*. New York Daily News, 2023-11-16.  
<https://www.nydailynews.com/2023/11/16/nikki-haley-big-sister-speech-even-anonymous-speech-must-be-protected-not-banned/>