

we asked

Expert Panel on Emergent Technology – proposal to trial Zavy

In May 2021 the Panel's advice was sought on a proposal to trial a technology called Zavy, a social media sentiment and analytics tool, which would allow for deeper insight into public sentiment towards New Zealand Police, based on analysis of the tone of social media posts and online content.

There are almost 2 million followers of Police social media channels and 4 million social media users in New Zealand at this time and Police are seeking to understand sentiment and performance of social media content on their own pages.

Zavy is a New Zealand-based company that provides brands with a strategic view of their performance across social and digital media. Sentiment measurement is a common form of social media analytics that has been used for many years to inform and help brands manage their social media presence.

EPET 21.3 Zavy Social Media sentiment and Analytics Tool

1. The Expert Panel has been asked to provide advice to the OCGG on a proposed trial of this tool.
2. The Panel had major reservations as to the proposal's compliance with the Necessity principle as set out in the New Technology Framework. Specifically, we were unconvinced that the use of the technology would deliver an identifiable public or policing benefit that would not otherwise accrue.
3. The proposal claims that "it is critical that Police has an understanding of ... sentiment (tone) of the commentary and engagement on the official New Zealand Police social media pages" and "Topics Kiwis are talking about publicly online that relate to Police." This understanding, we are told, "is important to help ensure the organization remains aware of, and responsive to, the values and reactions of the New Zealand public."
4. The Panel do not believe that comments on Facebook pages provide an accurate indication of the "values" of the New Zealand public. Commentators on social media are unlikely to provide a representative sample of the New Zealand population. The age, gender, ethnicity, social class etc of those who use (and importantly, comment on) social media would need to be controlled for if this sample is to have any methodological validity.
5. Equally importantly, it is widely recognized that users make deliberately provocative, inflammatory, sarcastic and 'edgy' posts on social media. The instantaneous nature of these platforms incentivize knee-jerk unreflective responses rather than deeper consideration. Great care must be taken if inferences about "values" are to be derived from such content. This is a well-documented concern about the use of sentiment analysis tools on social media content. As one academic paper has warned, "what appears as sentiment may in fact be its performance." The paper goes on to note that:

If users of social media like Twitter self-consciously self-construct, as some have suggested ... or if they conceal more than they reveal ... then whether sentiment analysis is taking place at all is debatable.¹

Another paper on social media sentiment analysis warns that "there are legitimate concerns about the reliability and accuracy of what is posted, whether posts are created by human actors or automated software, and about the intent of those posting material."²

¹ Helen Kennedy (2012) "Perspectives on Sentiment Analysis." *Journal of Broadcasting & Electronic Media* 56:4, 435-450

² Jeremy Prichard, Paul Watters, Tony Krone, Caroline Spiranovic & Helen Cockburn (2015) "Social Media Sentiment Analysis: A New Empirical Tool for Assessing Public Opinion on Crime?" *Current Issues in Criminal Justice* 27:2, 217-236

6. Consideration also needs to be given to cultural variation in modes of expression, and particularly, to the weight and value of silence in some cultures. Norms of communication and intensity of communication are important. Inferring acquiescence or indifference from silence would be highly problematic without giving due consideration to these different modes of engagement and expression.
7. The Zavy tool seems more suited to the objective noted at [5], “allow[ing] the Marketing and Brand team to understand sentiment and performance of Police’s social media content”. Our concern here is as to whether such market research is a sufficiently important objective to risk the erosion of public trust that is likely to accompany this trial.
8. The Panel notes the acknowledgment at [5], echoing a concern raised by the OPC, that “the tool currently drills down to an individual level ... which is not something, intuitively at least, that seems necessary for quantifying sentiment.” We share the view of the Police Chief Privacy Officer that limiting the ability to see private information is (at least) preferable, and endorse the OPC’s advice that, should NZ Police wish to pursue social media sentiment analysis, it looks for a “more privacy focused solution that does not enable Police to identify individual users.”
9. The Panel shares the OPC’s concern about the absence of a privacy policy on Zavy’s website. A Google search did reveal a privacy policy bearing their name, but it is out of date (referring to the previous Privacy Act.) We would hope that a company engaged in this kind of work would take seriously their obligations in this regard, and would wish to do so in a transparent fashion.
10. The Panel notes that the referral is of a trial of Zavy technology, but the specifics of this trial are not made clear. What would constitute a successful trial? How is success or failure to be measured? The Panel raises this as a general point for future consideration; proposed trials of new technologies will be impossible for us to fully evaluate without information about the evaluation criteria.
11. The Panel disagrees with the statement that this is likely to enhance public trust and confidence. While social media users may be taken to have consented to the general viewing of their public posts, it is unlikely that they will have considered the likelihood of police scrutinizing of their comments, particularly away from the Police’s own pages. The Panel considers that this is likely to be perceived negatively, potentially as a form of surveillance, and we echo the OPC’s concern that it may have a chilling effect on some people. While this may be justifiable if the rationale were sufficiently important and the methodology rigorous, the Panel does not consider either of those criteria to be met in this case.

we did

Police have taken the Panel's advice into consideration and have decided not to trial Zavy.