

Can Social Media Solve the Polling Problem?

Erin Kelly

Like so many elements of democracy, traditional public opinion polling has been overtaken by technology. The gap between technological adoption among consumers and technological innovation among pollsters has contributed to a parade of polling debacles whereby election-eve numbers and actual election results are wildly divergent. Advanced Symbolics CEO Erin Kelly describes how social media sampling enabled by artificial intelligence may be the solution.

Social media isn't just a broadcast technology. It's also a polling technology; a polling technology that is far superior to any other on the planet today, including the telephone. So why does the Canadian government continue to use telephone polling to find out what the citizenry is thinking? Why haven't public servants progressed to new tools for polling, as have the governments of France, the United Kingdom, and even municipal governments in Canada?

On March 15, the Privy Council Office—the prime minister's own department and the central agency of government—announced it had commissioned Forum Research to a \$312,000 one-year polling contract to conduct a weekly rolling survey of Canadians. PCO decided to limit itself to telephone surveys and as the *British Journal Nature* so aptly argued in an October 2016 article “*The Polling Crisis and How to Tell What People Really Think*”, telephone polling is so flawed that many governments have stopped using it. Consider the following:

- Fewer than 10 per cent of respondents will agree to participate in a telephone survey.

This creates what scientists call opt-in bias.

- Low opt-in rates force pollsters to over-sample minority demographics. Since opt-in bias tells us they are likely outliers to begin with, this leads to a very distorted view of public opinion.
- Merely posing a question to respondents creates question bias.
- Small sample sizes are not representative. The government's rolling survey will sample 500 people a week. That is not enough to tell us what people across the country, from every region, language and socio-economic group think. Not even close.
- Rolling or windowed surveys are controversial because they are trying to mimic longitudinal studies which survey the same people over a longer period of time. With new technology, there is no need for this mimicry. Real longitudinal studies are already possible, for a fraction of the cost the government paid for this substitute method.

The 2015 Canadian Federal election provides a good illustration of how telephone polls

can lead to wrong insights, and cause politicians and policymakers to make the wrong decisions.

Polls released a week after the viral release of a photo of the body of 3-year-old refugee Alan Kurdi washed ashore in Turkey showed a drop in Conservative support. The media immediately assumed that the drop in support was due to Kurdi's death, and “social media listening” on Twitter seemed to confirm this. Tweets about Kurdish and Syrian refugees increased 300-fold, as did criticism of the Harper government, which had been slow to let refugees into the country.

What Harper, the media and just about everyone else failed to recognize was that the fall in Conservative support had nothing to do with Alan Kurdi's death. The day before Kurdi drowned, Statistics Canada published a report saying that Canada had entered a technical recession. It was this report, and not the drowning, that caused the fall in Conservative support. People who voted Harper did so because they believed him to be a good steward of the economy. They were not concerned about his policies on refugees. Conservatives supported taking a cautious approach on immigration. And all those anti-Harper Tweets? This is a perfect example of why social media listening doesn't work (listening is not the same as sampling. Listening analyzes every Tweet, and Twitter is over-represented by certain demographic groups, so you need to correct for this through sampling).

Our artificial intelligence (AI) showed those anti-Harper Tweets came from committed supporters of the Liberals and NDP. They were never going to vote for Harper in the first place. How do we know this? Because the



A Leave mobile billboard makes its way through London before a TV debate between David Cameron, prime minister and head of the Remain side, and Nigel Farage, head of the Brexit insurgency in the June 2016 UK referendum on staying or quitting the EU. Wikimedia photo

AI is able to analyze the same people over the whole year. It knows where your true colours lie, and it doesn't need to ask.

It was the Harper government's over-reaction to the refugee issue that cost them the election. The Conservatives, in order to defend their refugee policy (which their supporters weren't asking them to defend in the first place), went overboard, promoting fears about Muslim immigration with the "barbaric cultural practices" tip line. This over-reaction turned the soft Conservative support away.

Because AI works 24 hours a day, it is able to see, up to the minute, when public opinion changes. It was able to distinguish between the Statistics Canada report and the photo of the Kurdish boy drowning. A phone poll taken a week after both events, and which cannot analyze the real leanings of the people taking the poll, was not able to make the distinction. With phone polling, if you are not asking the right questions, you have no chance of getting the right answers. Furthermore, the phone pollsters were questioning people when emotions were running high. The AI, by contrast, can analyze the

same people all year, and can predict how things will settle after immediate tragedies have passed.

We now have the ability to call up social media pages randomly the same way that we used to randomly dial telephone numbers. To be clear: We're not talking about web panels. Web panels are even worse than telephone polling for gauging public opinion. Samples garnered from web panels are neither randomized nor controlled—two necessary pre-conditions for interpreting public opinion. I am talking about randomized, controlled samples of social media users. This new technology is called Conditional Independence Coupling (CIC).

CIC (pronounced "Kick") was invented in 2012 and has been used successfully to predict over 100 elections and referendums around the world, including the recent Trump election in the United States, BREXIT and the Canadian Federal Election in 2015. Here's the kicker (pardon the pun): Since it was commercialized in 2015, CIC has never been wrong. Not once. It is being used by governments around the world, including

the United States, France and Britain. CIC does not require opt-in and it does not require the pollster to ask any questions (that's the beauty—the information is already there). Consider the following advantages of using randomized, controlled samples from social media:

- **Engagement Score.** CIC can tell us what percentage of the population is discussing a particular topic without being prompted. For example, if I call someone and ask "What do you think of the Islamophobia Motion (M-103)", the respondent will give me an answer even if, previous to this call, she didn't care about it at all. With CIC technology, we can gauge how important a topic is to Canadians, without asking them.
- **Large sample sizes.** We have a Canadian sample of 75,000 Canadians. Compare that with the sample of 500 the government's rolling survey is using. And we can survey that sample every hour if we wish to, and see how their opinion changes over time or even how it changed over the last year or five (we can go back in time with social media). We can even

predict how their opinion will evolve over time and what will change it. And we can do it for less than what the government paid for its telephone survey.

- **No need to “weight” under-sampled demographics.** With over 3,000 Aboriginals in our sample, we don’t have to “weight” the Aboriginals in our sample more heavily, we have enough of them in our sample to be confident we are measuring their opinions accurately.
- **No opt-in bias.** Privacy is meticulously protected, no names collected, only demographic information. The AI can map the demographics to the census to ensure a perfectly representative sample—no weighting required.
- **No question bias.** No questions are asked. Instead, the CIC algorithm is able to interpret what is being said to answer the questions posed to it—naturally and without biasing the sample.

“ *In the case of the 2016 BREXIT referendum, the CIC algorithm correctly predicted an EXIT at 52 per cent with a margin of error of +/- 1.2 percent. Our company was one of only 16 in the world to correctly predict an exit, and we did it using CIC.* ”

In the case of the 2016 BREXIT referendum, the CIC algorithm correctly predicted an EXIT at 52 per cent with a margin of error of +/- 1.2 percent. Our company was one of only 16 in the world to correctly predict an exit, and we did it using CIC. CIC’s AI read all social media topics over a yearlong period (not a weekly poll, a yearlong poll of the same individuals) and developed a prediction algorithm based on its

“ *Not everyone participates in social media and social media is over-represented by some demographic groups—like young people. But there are enough people on social media that we can get a representative sample of the population. Not everyone has a telephone landline either. Consider that, since 2010, there are more people with a social media account than with a telephone.* ”

analysis of how people reacted to different events and news items related to BREXIT. Using a sample of over 30,000 Britons from every region and walk of life, it was able to analyze how the opinions of Britons changed over the course of the year, which enabled it to correctly predict that the murder of MP Jo Cox a week before the referendum would change the outcome of the BREXIT vote. This would not have been achievable using telephone polling. Don’t blame the pollsters. It’s the technology that isn’t working anymore.

Back in August, 2016, when media commentators were poking fun at Donald Trump for alienating black and Hispanic voters, our company published a report showing the opposite—that support for Donald Trump was growing in exactly these demographics the pundits thought he was losing. Because the stereotype of the Trump voter was so negative (sexist, racist, lower-class), people were unwilling to admit to pollsters on the other end of the telephone that they were going to vote Trump. It was unpopular to admit such things in polite company. But online, our AI could see differently. It saw voters sharing anti-Clinton documentaries such as *Clinton Cash*, documentaries that were professionally produced by friends of the GOP, but which never saw mainstream play. As we now know, there was a whole stream of underground news taking place during the 2016 election that the intelligentsia knew nothing about.

Not everyone participates in social media and social media is over-

represented by some demographic groups—like young people. But there are enough people on social media that we can get a representative sample of the population. Not everyone has a telephone landline either. Consider that, since 2010, there are more people with a social media account than with a telephone. This is why we sample. Sampling ensures that we have enough people in each demographic to accurately assess public opinion. Contrast this with social media listening, where companies listen to everything that is said. This is bad, because, for example, 1/3 of all posts are made by bots. And many lobbyists are paid to Tweet, distorting results like we saw in the 2015 Canadian election. The CIC algorithm has dozens of algorithms that correct for this, ensuring we are reaching real, average people. Not bots, not influencers or pop stars. Average citizens.

The whole impetus for the populism movements taking shape around the world is that citizens do not feel they are being heard. It is still true that if you want to effect change in our society, you have to be well-heeled and organized. Policy is being made by special interest groups, lobbyists and big business. We want to change that. We want to make it possible for policy-makers to hear what citizens are saying so that policy reflects the needs of real people. That is the essence of good government—an engaged citizenry whose needs are being met. **P**

Erin Kelly is President and CEO of Advanced Symbolics, an Ottawa-based company that uses artificial intelligence (AI) for public opinion research.
erin.kelly@advancedsymbolics.com