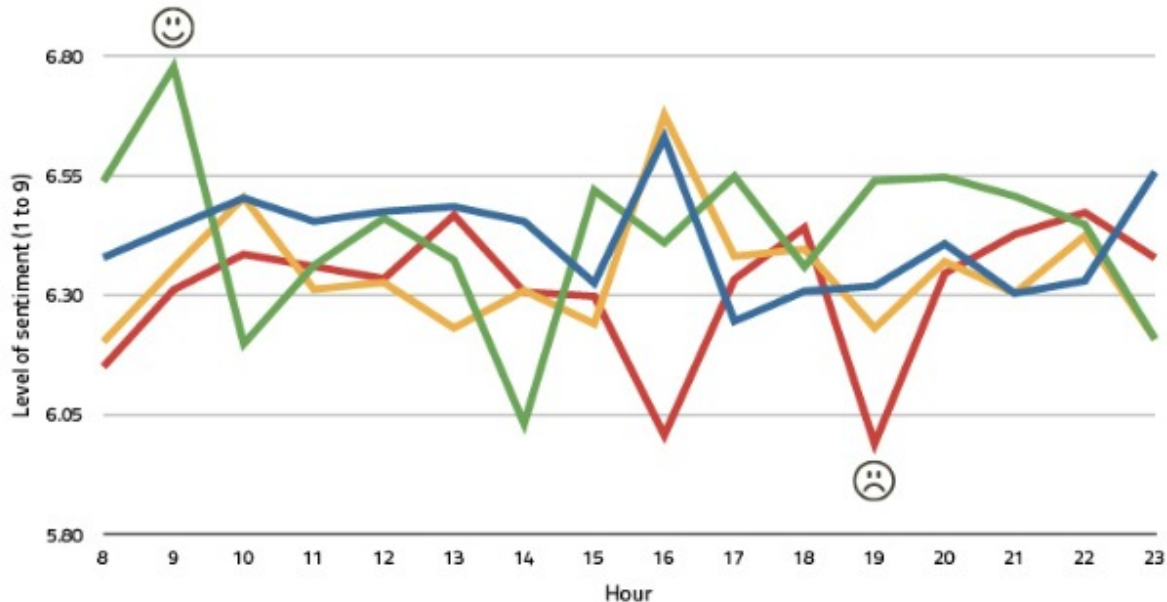# Businesses can no longer ignore social media sentiment analysis

blogs.lse.ac.uk/businessreview/2016/08/18/businesses-can-no-longer-ignore-social-media-sentiment-analysis/

In February 2015, Bridgewater Associates, a $165 billion hedge fund founded by billionaire investor Ray Dalio, contacted a number of artificial intelligence (AI) programmers. He offered them an opportunity to join David Ferrucci, who came to Bridgewater at the end of 2012 after leading the Semantic Analysis and Integration Department at IBM's T.J. Watson's Research Center for a number of years.

Their job would be to create AI algorithms that made predictions based on historical data and statistical probabilities. Part of this task would be predicting and minimising political risk by analysing social media data through algorithms.

Fast-forward a year and the investment seems to have paid off. Bloomberg notes that following the vote to leave the EU, Brevan Howard Asset Management – another firm which reportedly invested in artificial intelligence analysis of social network data – gained 1 per cent on its $16 billion macro fund. Hedge funds globally lost 1.6 per cent.

Polls and bookmarkers on the eve of the referendum suggested the vote was tight between the two campaigns. The markets betted en masse on a 'Remain' verdict. The pound surged before dropping to its lowest level since 1985 as the reality of a Leave vote took hold.

The failure to predict the outcome in the days and hours before the results demonstrates the increasing connection between Western political risk and business risk. Within this context the accuracy of political polling is increasingly business critical.

Opinion polls have been having a tough time in the last few years. Following the 2015 UK General Election, the polling industry launched a massive inquiry into why they had collectively failed to predict a Conservative majority. Its conclusions were long and drawn out, but in essence reduced the failure to one critical problem; people failed to report their true feelings and voting intentions when asked, skewing survey samples.

For many this 'selectivity bias' belies an inevitable and irrevocable problem with opinion polling. Polls suffer, and

have always suffered, because people often respond with the answers the pollsters want to hear, suppressing extreme or marginalised opinions.

This happens less online.

But analysing public sentiment online also has its flaws: chiefly, its own selectivity bias. In the U.K., about 33 million people are on Facebook, but the number actively posting opinions and views on Twitter (the easiest channel to analyse) is much, much lower. So low in fact that many have accused the platform of being an "unrepresentative urban liberal dreamland".

This conclusion is woefully inadequate. Although Twitter represents a large-scale bias towards the intelligentsia – often mockingly nicknamed 'the Twitterati' – it also represents the widest, most inclusive public conversation ever undertaken by mankind.

Every second, on average, around 6,000 tweets are posted, and whilst most of these are trivial and uninteresting, ignoring such a large data set would be madness in any science. The trick is to build representative samples within these tweets, just as you would with an opinion poll. After we have done this we bring in artificial intelligence.

Sentiment analysis is the critical advantage social media has over opinion polling. It allows analysts to gather the real opinion of voters/consumers rather than what they want telephone surveyors to know. But analysing what consists of positive or negative sentiment has always been challenging, leading most analysts to arbitrarily measure re-posts, likes and hashtag mentions.

This is where natural language processing comes in. Over the last few years, a virtual revolution has taken place in data analysts' ability to process and understand language in its own terms. At companies such as Google, Facebook, and Amazon, as well as at leading academic AI labs, researchers are attempting to finally solve how to teach machines to speak and understand language in the same terms as humans do.

The application of this technology to sentiment analysis online (such as the Deep Listen research project which launched this month) should be taken seriously in the battle to minimise political risk.

We enter a world where referendums and elections are becoming more and more common; where consumers' identities and allegiances are becoming more and more fluid; and where opinion polling is less and less reliable. We must approach such technology with enthusiasm and opportunism.

♣♣♣

Notes:

- *The post gives the views of its authors, not the position of LSE Business Review or the London School of Economics.*
- *Featured image credit: UK Twitter sentiment, Robin Hawkes, CC-BY-SA-2.0*
- *Before commenting, please read our Comment Policy*

Hugo Winn is a communications and policy consultant at Weber Shandwick. He specialises in technology and startup policy with a focus on content, privacy and internet-enabled services. He is founder of *Deep Listen*, a project that will use natural language processing to recognise the nuances of positive and negative statements, in an attempt to correctly predict the next leader of the Labour Party, three days before it is announced in September this year.

-