

Human integration the key to accurate social media data

By JP Kloppers

28 Oct 2016

The digital age has fundamentally altered how we communicate - particularly via text. Be it a Tweet, online comment or Facebook post, we're constrained not only by the limits of language's written form, but often by a character limit. Seeking to put across as much information in as few words as possible, we've changed how language works.

We can all attest: communicating without non-verbal cues such as expression, tone of voice, eye contact and body language can be tricky. A nod, a shrug, a smile, a sigh – take these away and it becomes far harder to put a short message across clearly. Were this not the case, there'd be no need for emojis – the tiny happy/sad/laughing faces we use to make sure our messages' meaning hits home.

These virtual expressions usually confirm sentiment, but can also alter it. The addition of an angry face, for example, quickly takes the feeling behind 'So happy that my flight was delayed!' from possible relief to genuine annoyance. A smiling face transmutes the tone of 'I can't believe you did that!' from likely irritation to effusive appreciation.

An all-seeing algorithm?

As in other areas of our lives, we have tried to digitise and automate: if conversations are happening via machines, why not get machines to analyse what's going on in those interactions, too? This has been the logic behind the algorithms developed to try to analyse the meaning and sentiment that drives online chatter. The only problem is: machines often get it wrong.

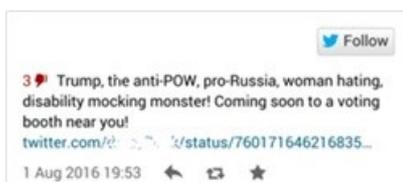
As advanced as computing power has become, it remains overly optimistic to ask a machine to comprehend all the nuances of human communication. Language is, after all, far more than mere syntax – it is a culturally-embedded mode of expression that relies on shared (human) experiences.

Consider the following examples:

- 'Thanks so much!'
- 'Great! I can't believe this is happening!'
- 'Yet another great piece of news about the government!'

- 'I absolutely hate that!'

Although simple-seeming at first, all of the above have possible dual meanings. The first sense is at face value: genuine appreciation, pleasant surprise, delight and disdain, respectively.



Alternative meanings, however, stem from sarcastic readings – an acerbic thanks, two expressions of irritation, and a joking remark about dislike. Consider the Tweet (left).

While it may be obvious to us that the speaker above is being sarcastic about Donald Trump, a machine simply wouldn't get the joke. Going on textual content alone, it would probably rate the sentiment behind the message as positive. Understanding the real meaning relies on insight into how we humans communicate – which is, often, anything but direct.

Man versus machine

Due to the complexity of the human language, the solution to understanding it is quite simple – get humans, not machines, to do the analysis. The only approach that seems to get it right include people, or in our instance a global crowd. Only three companies in the world currently use this unique data processing on scale approach, namely: Crowd Flower, Amazon Mechanical Turk and [BrandsEye](#).

One of the traditionally limiting factors that has impeded the uptake of [crowd](#) driven data processing has been the generalised approach to the workflows within certain crowd sourcing platforms.

The solution, however, is simple, an end-to-end solution where human verified data is integrated into your analytics journey from start to finish.

Helping machines learn

Through this unique approach, the crowd is able to both confirm (or contradict) machine understanding, and allow human insight to help ‘teach’ machines about the nuances and cultural shifts of language.

After all, sarcasm isn’t the only area where machines struggle – local references, slang, ever-changing cultural memes and words with in-built double meanings also pose a challenge.

Consider companies’ taste for simple, memorable names. Anything from Apple to Amazon, Blackberry, Windows, Checkers, Oracle and others – machines are asked differentiate brands from their namesakes. This is not always easy, and often requires manual coding by those behind the scenes.

Next consider slang. Terms such as ‘hyped’ (excited), ‘knock’ (insult/deride), ‘on point’ (perform well/meet standards), ‘slay’ (achieve something amazing) and ‘shout out’ (acknowledge/greet) are all widely used, particularly among the youth, though they came into the vocabulary only recently. However, to the uninitiated, their meanings are far from obvious. Consider how an algorithm might struggle to classify the Tweet (right), which uses both slang and sarcasm.



Acronyms are also problematic for machines. Having expanded from the simple ‘LOL’ (laugh out loud), they now encompass everything from ‘IMO’ (in my opinion), ‘JSYK’ (just so you know), ‘ICYMI’ (in case you missed it) and ‘IRL’ (in real life), to ‘NSFW’ (not safe for work) and the much-derided ‘YOLO’ (you only live once). Of course, this is just a taste of what’s out there, and only what’s within the English-speaking realm.

Expecting a machine to correctly decode the sentiment behind a post filled with pop-culture slang and self-referential hashtags isn’t just unrealistic, it also approaches the nature of language and communication incorrectly.

People aren’t machines – swapping straightforward statements and observations – but are instead complex, culturally embedded entities who change the ways in which they share information with one another, and adopt new lenses through which to see the world. They use current events as rallying points, invent new words and subvert the traditional meanings of older ones.

The best thing to do? Ask real people – not machines – to make sense of what they’re saying. Using an integrated crowd approach ensures you have truly accurate social data that provides you with actionable insights that you can use to make informed business decisions.

ABOUT JP KLOPPERS

CEO at BrandsEye

- Do you know how people really feel? - 6 Jul 2017
- Social analysis shows majority of Americans against Trump's immigration order - 2 Mar 2017
- Data analytics compares social media chatter surrounding Pzza Hut in UK vs SA - 14 Dec 2016
- Social media sentiment shows why customer expectations matter - 8 Dec 2016

▫ When fame is no longer enough: Brand ambassadors in the age of social media - 5 Dec 2016

[View my profile and articles...](#)

For more, visit: <https://www.bizcommunity.com>