

Word clouds – the bête noir of the text analytics industry!

We're often asked about them – do we do them? how useful are they? etc. With data visualisation becoming so popular, and easy-to-use word cloud software being so readily available, you can understand why they have caught on. However, one of the most popular and 'cool' solutions for creating word clouds, "Wordle", describes itself as "a toy". You really have to wonder how worthwhile these word clouds really are as an analytical tool!

Rather than simply be disparaging about word clouds, let me explain why here at Feedback Ferret, we don't hold them in high regard.

In our view, trying to transform a piece of text into a visualisation of what the most important themes are cannot simply be done by boiling it down to a group of individual words and making the most-used words larger than the others. This method takes those individual words totally out of context and ultimately, words have no meaning without context. And without context, what insight can **really** be gleaned from a random sample of words?

Here's where word clouds fall down:

They do not cater for negatives – how often have you seen the word "don't" in a word cloud?

- They do not cater for misspellings.
- They do not cater for different words or phrases meaning the same thing (bad, rubbish, poor, disappointing, terrible).
- They do not consider **strength** of sentiment or emotion.
- They do not consider false positives or false negatives (eg "I would not trust my car service to anywhere else")
- They do not consider the meaning of ambiguous words ("light" could mean lightweight, pale, lamp, casual, gentle, a traffic signal etc)
- They do not take into account figures of speech (eg "crying wolf" or "bull in a china shop" or "white as a sheet")

Word clouds simply make the words that appear the most often the biggest. Which is just plain silly. ▶

A word cloud cannot group these together to provide insight into how good or poor something is because there are so many variations and different words pertaining to poor (or good) in every language.

A friend of mine recently did one of those Facebook word clouds. It took everything he has ever posted and placed it in a cloud. The biggest word was Susie, his wife's name. Well, of course Susie is the most important thing in his life. He knew that. Susie knew that. So what did he glean from this exercise? Absolutely nothing.

As the word 'cloud' suggests, they have no substance. They are, quite simply, just pure fluff. Or as Jacob Harris, Senior Software Architect at the New York Times describes them – “filler visualization” in his article ‘Word Clouds Considered Harmful’.

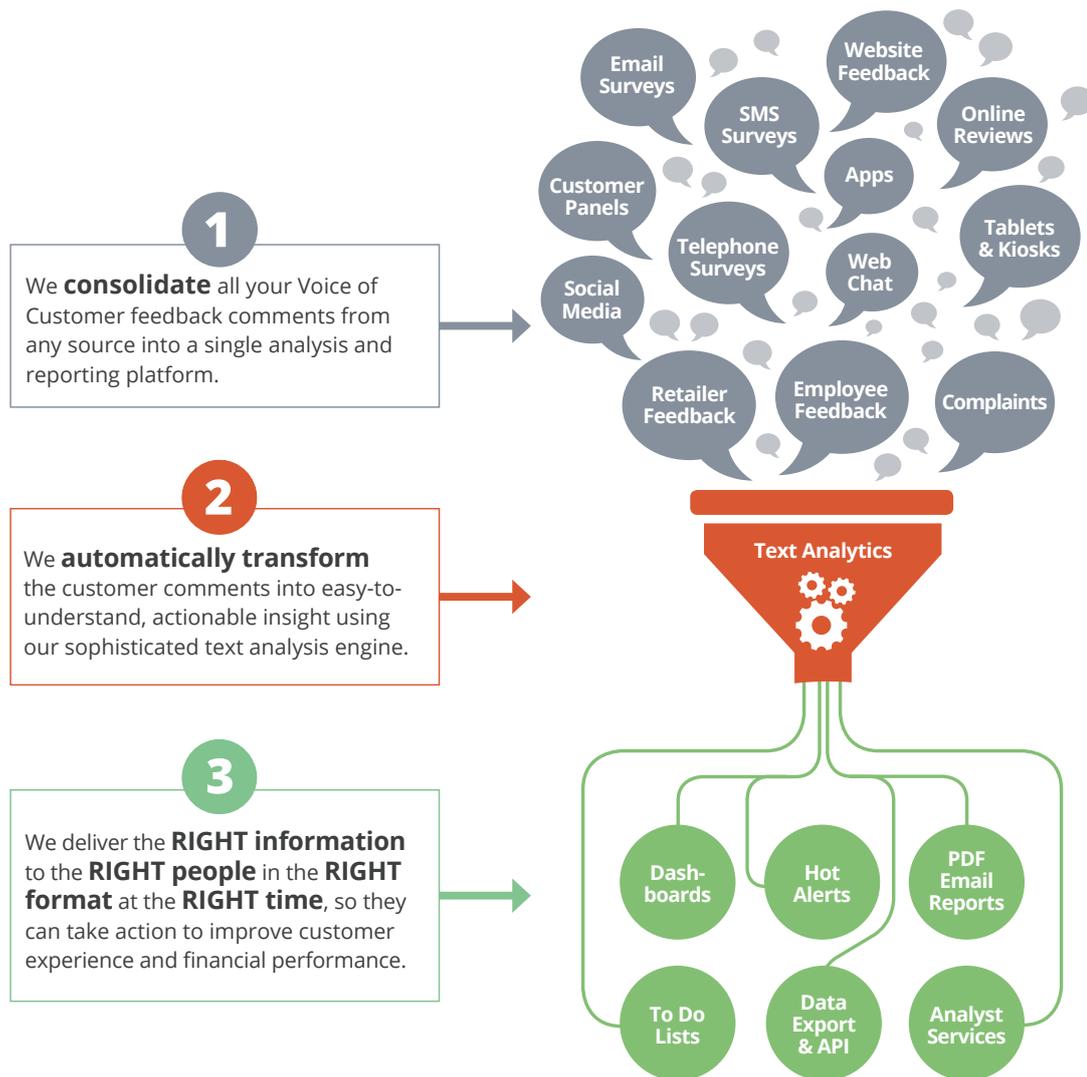
That's all the hot air we are going to put out about word clouds. And in case you were in any doubt – no, we don't ever show word clouds!



**Author – Piers Alington, Managing Director,
Feedback Ferret**

Feedback Ferret – Experts in Customer Feedback

Using text analytics, Feedback Ferret transforms customer feedback comments into actionable insight. We make it easy for you to improve your customer experience.



All our technology and services have been developed by, and are operated by, Feedback Ferret Ltd. The technology is our own intellectual property, designed specifically to address the needs of extracting insight from customer feedback.

For more information visit:
www.feedbackferret.com