

# Getting Started in Text Analytics

Fiona Moss

The questions to ask

# Getting Started in Text Analytics

Fiona Moss

As the number of text analytics tools have proliferated, the market has moved past early concerns about accuracy and now sits in a place where picking the right tool is more about ensuring the outputs will fit your needs.

This is no small challenge, not least because jargon and terminology abound in this market. From semantic engines that identify word patterns then use rules-based approaches to build categories, to probabilistic/statistics-based approaches that use frequencies and co-occurrence of words to derive results, it can be difficult to know exactly what you are getting.

We outline below a few key questions which may help.

## The key questions:

### **1. Is your focus on exploring the data or categorising and quantifying it?**

Some tools offer sophisticated outputs and visuals for exploring the content of the data, but little by way of quantification. Other tools' strengths lie in quantification, while some can address both. Knowing where your focus lies will help the selection of the tool.

### **2. How much data do you have?**

A starting point of half a million comments can justify a very different investment to a starting point of 5,000 comments, meaning tools with a lengthier set up process are typically suitable only for higher volumes.

It is worth also bearing in mind that many tools have minimum base size requirements for accurate set up.

### **3. How much consistency will you need in the analysis over time or between projects?**

There are two broad options here: the first allows a categorisation to be fixed and saved post-set up, ensuring completely consistent processing wave-on-wave or between different sources. The second option involves the tool 'updating' itself to adjust for different content between analysis runs, making results less comparable.

### **4. How much transparency and control over the categorisation do you need?**

Some tools offer complete transparency over why each comment is categorised in the way it is – and the means to adjust the categorisation if necessary; other tools can be a little more black box about the way they categorise comments, and harder to update where the categorisation is incorrect.

# GETTING STARTED IN TEXT ANALYTICS



May 2016

Fiona Moss

If your business is likely to need evidence of how comments are categorised in order to buy into text analytics as a process, then it is important to consider this upfront.

## 5. Is there a lot of noise in your data (for example, do your social media comments contain a lot of irrelevant posts)?

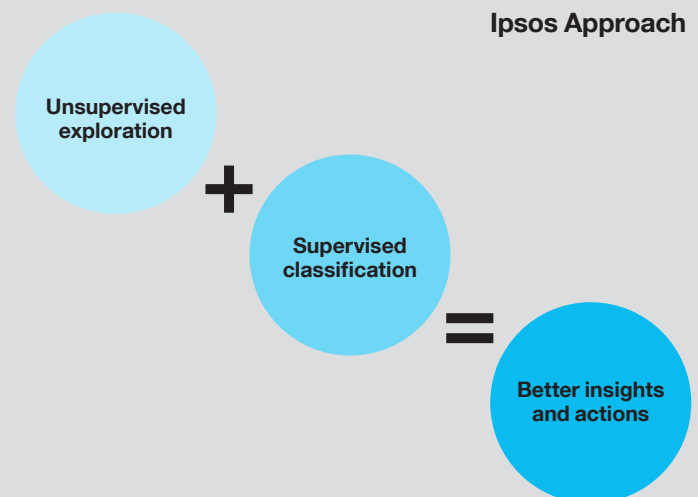
Some tools now allow relevance filtering prior to text analytics proper; others can categorise in quite a specific way, meaning that any nonsense/irrelevant comments are left out of the model. However, in some cases a lot of noise in the data can confuse tools, leading to spurious categorisations.

## 6. How much granularity will you require?

Granular results rely on there being sufficient detail in the comments under analysis. It is therefore worth checking the depth of information relevant there.

Most tools will build granularity to match the level of detail available in the comments (assuming sufficient base sizes). Many will allow the analyst to force further granularity and/or search for specific themes of interest, so this should be mentioned if you think you may require it.

Each approach has pros and cons depending on use cases. Ipsos uses a small portfolio of tools that enables us to use the right tools in the right way for any given situation. We use semantic engines and rules-based approaches to extract sentiment and themes from text and to create structured/hierarchical categorisation of the data, but can also leverage more exploratory and probabilistic models to identify the key relationships in any given dataset.



# GETTING STARTED IN TEXT ANALYTICS

September 2016

---

Fiona Moss

**Fiona Moss** is an Associate Director in Ipsos Loyalty's Global Client Solutions team.

**Ipsos Loyalty is the global leader in customer experience, satisfaction and loyalty research with over 1,000 dedicated professionals located in over 40 countries around the world. Our creative solutions build strong relationships which lead to better results for our clients. This has made us the trusted advisor to the world's leading businesses on all matters relating to measuring, modelling, and managing customer and employee relationships.**

This *Ipsos Views* paper is produced by the **Ipsos Knowledge Centre.**

**[www.ipsos.com](http://www.ipsos.com)  
@\_Ipsos  
IKC@ipsos.com**

## GAME CHANGERS

<< Game Changers >> is the **Ipsos** signature. At **Ipsos** we are passionately curious about people, markets, brands and society. We make our changing world easier and faster to navigate and inspire clients to make smarter decisions. We deliver with security, simplicity, speed and substance. We are Game Changers.

GAME CHANGERS

