

March 2019

---

# IPSOS VIEWS

---

## High Hopes: Tips for ensuring successful text analytics

By Fiona Moss



The popularity of text analytics has grown exponentially. Now, far from being a niche analytics offering, it is an established component of many market research programmes, particularly when it comes to analysing the voice of the customer. As its popularity has increased, so too have expectations of what it can deliver. Among the success stories, there are inevitably some disappointments. In an industry in constant development, what can be done to ensure success?

## Recognise text analytics providers fuel expectations

It is no surprise that expectations of text analytics (TA) are high. TA providers themselves have been evolving and developing new products furiously over the past few years. The natural result of this is perpetuated hype, excitement and announcements that the next thing is better/faster/more accurate than the last. Consequently, the bar is set high for TA's performance, and there is little to no discussion of its potential limitations.

Alongside the hype, the TA industry can feel impenetrable. Different techniques abound, each with their own jargon, and sometimes just in the name. Distributional semantics, anyone? How about Latent Dirichlet Allocation...? Moreover, much of this analysis relies on specialist skills in artificial

intelligence, statistics, data science, linguistics, programming – or all of the above. And, given the complexity of the field, it is highly recommended to keep these specialists involved.

For those on the outside though, this means that many TA techniques, if not black box, certainly do not lend themselves to an easy understanding of how they work. While it is not an issue itself, if you don't actually have to carry out the analysis, this does not make it any easier to unpick an overclaim from reality.

This can be compounded by confusion around accuracy statistics; what exactly is meant by accuracy? Do we look inside topics to see how many comments have been correctly assigned to them? Or do we look across the entire dataset at how many comments should be assigned to existing topics but are not? Do we consider that perhaps not all salient topics have been identified? And how can accuracy figures be calculated when, by the very nature of many TA datasets, manually working through the entire dataset is impossible?

Accuracy statistics do of course exist, and the hype associated with TA is often well justified, but asking questions about the potential limitations of any technique is often a useful way to get under the skin of an approach and decide if what you are looking at is right for you.

## Define your objectives first and then match the tool to them

TA is a powerful tool for understanding text data easily and efficiently. How else would it be possible to understand the contents of half a million comments in a few hours? Or quickly and cost-effectively know what is said in several million comments, in multiple languages and in real-time, over the course of a year?

TA's power comes from using the right tool in the right way. For this reason, rather than picking one and sticking with it, Ipsos uses a small portfolio of tools and techniques that can be matched to the appropriate project.

Much of this matching will depend on the objectives of the analysis. It is therefore essential for success to define them upfront. Is the goal to understand the main concepts in the data and how these link together? For example, to



**It is essential for success to define the objectives of the analysis upfront.**

establish the mental networks of consumer perceptions of a brand? Or is it to sort comments into phrase-based categories that explain exactly what should be improved in the customer experience? Or is it, perhaps, to derive and track prevailing sentiment on social media comment?

Being aware that different tools have strengths in different areas – and deliberately picking tools that are aligned with project goals – is key to helping TA meet expectations.

**TA is not just about matching approaches to objectives. It is also essential to ensure that the project is suitable in the first instance. As a rule of thumb, this means thinking about scalability.**

Many TA tools have minimum base sizes below which machine-learning, natural language processing or statistical techniques cannot be accurate. Other tools may be able to cope with smaller base sizes, but in a way that is not cost-effective compared to more traditional approaches, such as standard coding. No TA approach will deliver more insight on 30 verbatims than can be established from reading them. Keeping comment volume in mind is therefore key when balancing TA's expectations against delivery.

### Make sure the data is fit for purpose

All that TA can do is reflect the contents of the data that it is based on. Even with the right tool in place and a huge dataset to work with, if there is nothing of interest said in the dataset, TA will not be able to find anything interesting.

This is particularly the case when thinking about social media analysis. If no-one is talking about a particular component of your offer, then TA will not be able to deliver any insight on it – and this is before we think about whether the sample of consumers talking online are representative of the wider consumer group of interest.

Equally, TA might identify that customer service is a strong part of your offer through your customer experience survey. But if no-one has

provided any detail about what exactly it was that made your customer service so good, TA will not be able to tell you this.

**It sounds tautologous to say that TA cannot invent what is not there, but it is easy to get caught up in the hype and excitement of new technologies and forget that the input data still has to support the objectives of the analysis too.**

With this in mind, Ipsos has a tried-and-tested framework of the best questions to ask to support successful TA in large-scale studies, including how to prompt for more detail to deliver granular, actionable results, while not asking too much of the respondent.



## Remember, text analytics does not have to stand alone

There is a temptation to consider TA results in isolation. This is not surprising as it is an analysis technique that often involves its own separate set-up, timings and visualisations in any given project. However, to consider it alone is sometimes too reductive, and means we miss out on the value of setting TA results in context.

There are several ways to ensure that TA is not isolated. The first is to tie the TA of open-ended questions into more quantitative measures from within the same traditional survey. At its most basic, this can mean simply cutting the TA results by different subgroups or KPI scores. More elaborate approaches also exist and, for example, can explore which TA topics have the biggest negative or positive pull on a KPI measure in order to help action planning. Thorough investigation and interpretation of TA results, combined with further analytics, are often a straightforward way to extract maximum insight from them.

The second approach is to establish a clear research ecosystem that sets the TA project against other relevant projects and enables links between them to be found. For example, this could translate into recognising that social media analysis, a customer survey and an employee survey all contain text data that delivers insight on the highs and lows of the customer experience.

Identifying these parallels would allow a comparable TA model to be generated across all three datasets from the outset, meaning a holistic picture of the



**To consider text analytics alone is sometimes too reductive, and means we miss out on the value of setting text analytics results in context.**

customer experience can be developed, including exploration of where the differing perspectives are aligned or disagree. Moreover, this approach can also be adopted to pull in non-survey metrics such as complaints data or contact centre comment.

Ipsos has developed an ecosystem approach that helps establish the links between different research projects and other data sources by identifying their core purpose and frequency of collection. The resulting map of different projects and data sources can be used not just to maximise the value from TA, but also to see how different data sources can inform one another more generally.



“

**It is only through the ambition and drive of TA's users that the market will continue to evolve and innovate, moving ever closer to tools that can deliver on our most exacting expectations.**



## In summary

My argument is not to be less demanding or ambitious in what TA can deliver. It is only through the ambition and drive of TA's users that the market will continue to evolve and innovate, moving ever closer to tools that can deliver on our most exacting expectations. However, there are other things we can do to ensure TA's success, as outlined in our 5 tips below.

## TIPS FOR ENSURING TA SUCCESS

- 1** Challenging TA providers to give a realistic view of what is and isn't possible
- 2** Defining upfront what the TA should deliver so that the right tool and approach can be put in place
- 3** Putting forward larger datasets for analysis, because this is the best way to get good-quality, cost-effective results
- 4** Making sure that data is actually relevant to the research objectives in terms of topics and degree of detail
- 5** Remembering that significant value can be derived from TA when it is considered within the context of other data sources

TA is a useful, exciting and growing field. But for the moment, we urge a note of cautious realism: TA tools are not perfect, and they are just machines. Yes, they are incredibly sophisticated, but until we create machines that can replicate human intelligence, TA tools will never deal with text in the same way that a human can.

For this reason alone, it is not just desirable but also *advisable* to keep an analyst at the heart of the process. And in many ways, that seems the most customer-centric thing to do, anyway.

*Further reading:*

*Getting started in text analytics:* <https://www.ipsos.com/sites/default/files/2016-09/Getting-Started-in-Text-Analytics.pdf>

*A guide to text analytics:* [https://www.ipsos.com/sites/default/files/publication/1970-01/Loyalty\\_Text\\_Analytics\\_Guide\\_PDF\\_Oct2015.pdf](https://www.ipsos.com/sites/default/files/publication/1970-01/Loyalty_Text_Analytics_Guide_PDF_Oct2015.pdf)

*Don't kill the analyst just yet:* [https://www.ipsos.com/sites/default/files/2016-07/Dont\\_Kill\\_Analyst-Feb2016.pdf](https://www.ipsos.com/sites/default/files/2016-07/Dont_Kill_Analyst-Feb2016.pdf)

*Decisions, decisions:* [https://www.ipsos.com/sites/default/files/2017-03/IpsosLoyalty\\_POV\\_DecisionsDecisions.pdf](https://www.ipsos.com/sites/default/files/2017-03/IpsosLoyalty_POV_DecisionsDecisions.pdf)

**Fiona Moss** works in Ipsos' Global Customer Experience Analytics team, where she leads the text analytics offer. Her focus is on developing best practice text analytics techniques that deliver accurate and actionable findings to her clients as swiftly as possible.

**Fiona Moss** Global Research Director, Customer Experience, Ipsos

Ipsos is the global leader in designing, measuring and delivering value from Customer Experience programmes. We help organisations retain customers and recover those at risk, grow share of spend, increase advocacy and drive up operational efficiency; to 'deliver a Return on Customer Experience Investment (ROCXI)'. Based across 50 countries, our expert CX teams help organisations at all stages of CX measurement and management, using a unique blend of research, technology, analytics and advisory solutions.

**www.ipsos.com**  
**@ipsos**

The **Ipsos Views** white papers are produced by the **Ipsos Knowledge Centre**.

#### **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**

