Getting started with text analytics

Take your CX to a whole new level

an eBook by

Kapiche

Introduction 

Customers are constantly reaching out to share their opinions about products and services: what they love, what they hate and the changes they think an organization should make to do better. These interactions—in the form of product reviews, customer feedback survey responses and social media—generate a goldmine of unstructured data for organizations. The problem is that extracting actionable insights from all of that raw information can be an enormous challenge.

> ***Text analytics is the process of extracting meaning from written communication.***
>
> ***In a customer experience context, text analytics refers to analyzing text that was written by customers about your brand, product or service. This can be in the form of NPS, CSAT, CES, customer feedback surveys or product reviews. It could even mean tweets or social media interactions. Using text analytics, companies are able to find patterns and topics of interest, and then take practical action based on what they learn.***

Many organizations already use text analytics to make sense of large amounts of unstructured customer feedback, but few leverage it to its fullest potential. Applying text analytics to your CX data is not complicated and can have a huge payoffs for your organization such as being able to find the specific "problem areas" of your product, reach out to frustrated customers quickly and more meaningfully, and better understanding how changes to your product directly affect customer sentiment.

It might sound complicated, but there are few hurdles to jump over when convincing your company to start making use of text analytics. First, you help them to recognize the time and budget savings. Then, you work on acquiring the proper tools to achieve your organization's goals. Luckily, we've pulled together an eBook full of everything you need to know to make an informed purchasing decision and get the most out of your text analytics program. Here's what you can expect:

Part 1: All about text analytics
– What's so great about text analytics?
– What can text analytics do for you?
– How does it work?
– Adding in some emotion with sentiment analysis

Part 2: Making it work for your business
– Who can make it work?
– What do you need?
– What should you avoid?

Part 3: The Kapiche approach
– Powering customer-centric companies

PART ONE

All about text analytics

What's so great about text analytics?

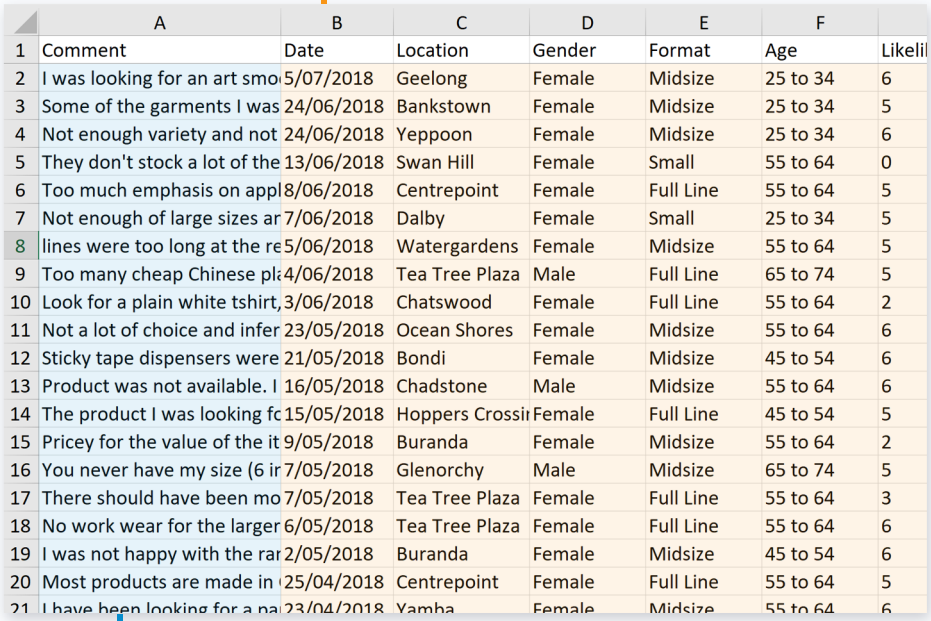
Data is at the heart of every business. Whenever an organization makes a decision, they should be basing that decision on the data they have on hand. So, what data does your organization have available and what are you using? Are you making the best decisions for your business? Common sources of data typically include: responses from customer surveys, tagging metrics from your helpdesk, and anecdotes or reviews from customer interactions. In this section, we'll take a look at the most influential, and most overlooked, type of data: unstructured data.

There are many different types of data available to an organization: customer data, transactional data and employee data. All of this data is broadly split into two categories: structured and unstructured. Structured data includes customer data such as age, location and gender and spending habits. These are hard, quantitative numbers that don't change subjectively through the customer's journey with your brand. In market research, this also includes limited questions like multiple choice and rated responses.

However, it is estimated that structured data only accounts for 20% of all available data. That means that the other 80% is out there somewhere, waiting to be understood. This unstructured data is where the opportunity is for businesses that want to understand their customers better than the competition does.

STRUCTURED DATA

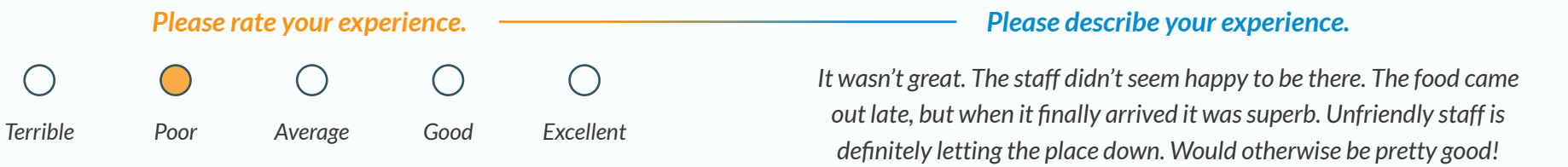
Data that resides in fixed, predefined fields. Most basic customer data falls under this category – usually collected via forms, surveys, etc where the respondant has fixed options to answer with.



	A	B	C	D	E	F	
1	Comment	Date	Location	Gender	Format	Age	Likeli
2	I was looking for an art smo	5/07/2018	Geelong	Female	Midsize	25 to 34	6
3	Some of the garments I was	24/06/2018	Bankstown	Female	Midsize	25 to 34	5
4	Not enough variety and not	24/06/2018	Yeppoon	Female	Midsize	25 to 34	6
5	They don't stock a lot of the	13/06/2018	Swan Hill	Female	Small	55 to 64	0
6	Too much emphasis on appl	8/06/2018	Centrepoint	Female	Full Line	55 to 64	5
7	Not enough of large sizes ar	7/06/2018	Dalby	Female	Small	25 to 34	5
8	lines were too long at the re	5/06/2018	Watergardens	Female	Midsize	55 to 64	5
9	Too many cheap Chinese pl	4/06/2018	Tea Tree Plaza	Male	Full Line	65 to 74	5
10	Look for a plain white tshirt,	3/06/2018	Chatswood	Female	Full Line	55 to 64	2
11	Not a lot of choice and infer	23/05/2018	Ocean Shores	Female	Midsize	55 to 64	6
12	Sticky tape dispensers were	21/05/2018	Bondi	Female	Midsize	45 to 54	6
13	Product was not available. I	16/05/2018	Chadstone	Male	Midsize	55 to 64	6
14	The product I was looking fo	15/05/2018	Hoppers Crossi	Female	Full Line	45 to 54	5
15	Pricey for the value of the it	9/05/2018	Buranda	Female	Midsize	55 to 64	2
16	You never have my size (6 ir	7/05/2018	Glenorchy	Male	Midsize	65 to 74	5
17	There should have been mo	7/05/2018	Tea Tree Plaza	Female	Full Line	55 to 64	3
18	No work wear for the larger	6/05/2018	Tea Tree Plaza	Female	Full Line	55 to 64	6
19	I was not happy with the rar	2/05/2018	Buranda	Female	Midsize	45 to 54	6
20	Most products are made in	25/04/2018	Centrepoint	Female	Full Line	55 to 64	5
21	I have been looking for a pa	23/04/2018	Yamba	Female	Midsize	55 to 64	6

UNSTRUCTURED DATA

Data that isn't easily categorized into a database. This eBook will primarily focus on text data, but unstructured data can come in the form of images, videos, audio, and more. This data is extremely hard to analyze due to its non-structured nature.



Understanding unstructured data

In unstructured data, information is not limited to clear, straightforward patterns. It usually takes the form of written information and can come from places like customer feedback survey responses, social media platforms and customer reviews. Without the limitations associated with structured data, the amount of usable information in unstructured data can vary greatly from one piece to the next, especially without a way to properly correlate it.

Take, for instance, the data that you receive in the 'comments' field of customer surveys. The responses you get could be about anything from the cost or quality of the product to the timeliness of delivery. The answers could be a positive response to your brand, a negative one or even a mixture of the two. Without analyzing the comment and its context, there is no simple way to draw a conclusion on what the customer is trying to say. These comments can provide much more detailed insights into the customer's needs, and they're unfortunately some of the hardest insights to get a handle on.

What can text analytics do for you?

Because of its lack of constraints, unstructured data can be a goldmine for businesses, especially in the CX space. This type of data provides you with a chance for you to immerse yourself in your customer's thoughts and feelings towards your product or service, and to truly understand what they think about their experience with your product or service in their own words.

Contrast this with structured data survey responses, where the customer is asked to answer a specific individual question or set of questions within defined guidelines like a rating scale. While this can be great for obtaining information on a specific topic, it prevents customers from giving insights on issues they are having if they happen outside of the scope of the survey.

For example, if you send out a customer effort score survey asking a customer "How easy was it for you to find an answer to your question today?" and ask them to rank their response as a 0-10 on a scale, you'll get a ton of quantitative data. This is helpful, to an extent—it helps you understand

if your customer service is easy to use, but it doesn't help you to understand what you could do to change it, or why you were given that number in the first place. The 0-10 scale is a great example of structured data.

If you added a second question after the scale that asked the customer what you could do to better your score by just one point, this would provide you with a bunch of unstructured data. You'd get a ton of great information, but it would be hard and time-consuming to find correlated data between each individual response.

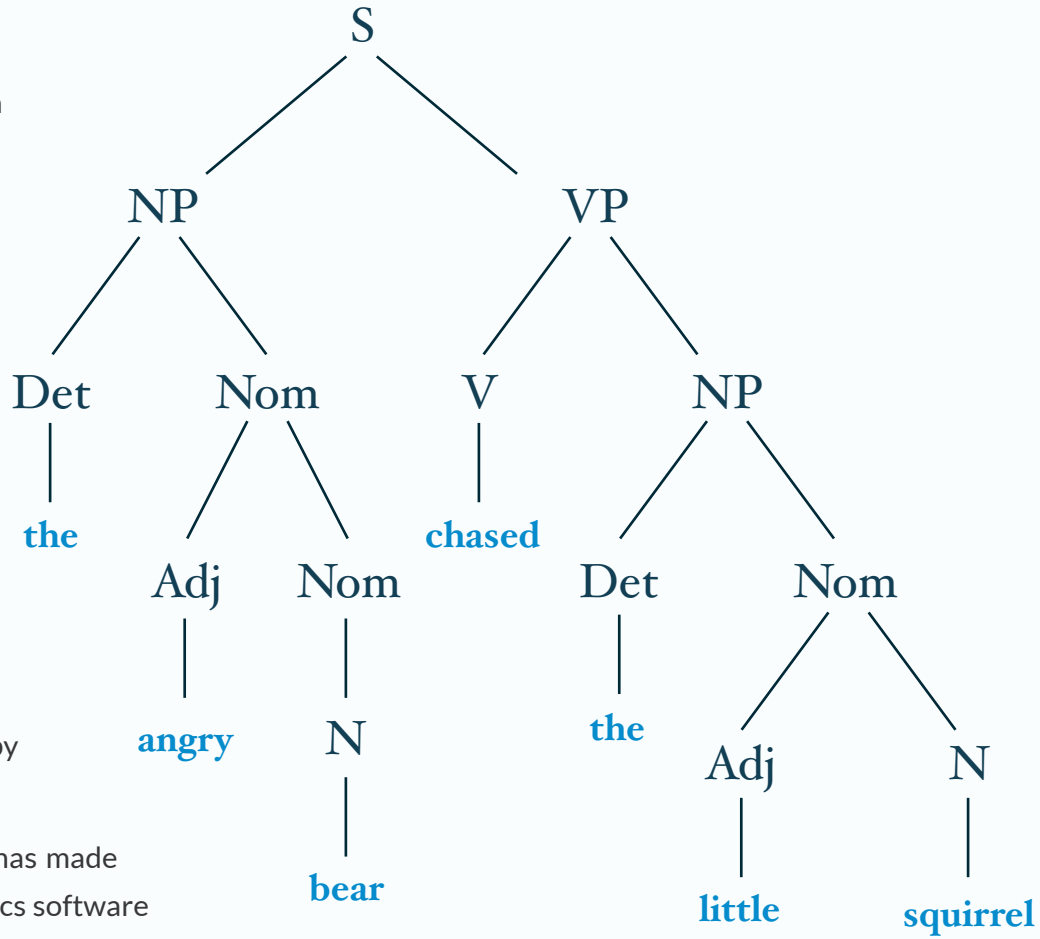
It is important to note that structured and unstructured data aren't mutually exclusive and ideally shouldn't be analyzed separate to each other. The best outcomes and interpretations come at the intersection of these two. By overlaying the unstructured with the structured you gain deeper insights about specific segments of your customer base. For example, link customer demographic data with survey responses to drill down on certain customer segments. You could do this querying what women on the east coast in the age bracket '25-34' think about your product or service versus those in the same age bracket on the west coast. Or, in the example above, you can use the structured data from the CES ranking score in combination with the qualitative data left in the comments section to fully understand what would make your customers' lives easier.

Analyzing unstructured data

Given that the benefits are so great, why isn't every business investing in analyzing their unstructured customer feedback data? We believe there are actually a few different reasons:

- Historically, analyzing this data has required a labor-intensive and costly manual categorization of text responses—most companies don't have a member of their team that they could take off of their regular job duties to review survey responses instead for days at a time.
- The larger the dataset, the more time it will take to sort through it. Having a human look through a large amount of responses would take ages and be extremely costly.
- Analyzing large amounts of data manually also introduces discrepancies due to human error and bias: what one person might consider a key theme of the responses might be totally overlooked by another, leaving the results subjective and often unrepeatable.

This is where the introduction of Natural Language Processing (NLP) has made such a difference to the world of analytics. With its creation, text analytics software platforms were able to quickly, accurately and repeatedly analyze large chunks of unstructured data in next to no time—a far cry from how long it would have taken someone to review it manually. Because of the increase of interest in the technology, using NLP has also become more affordable, making it much more accessible for all market research and both internal and external experience teams. In the next section, we'll go into more detail on what NLP is, and how CX teams can harness its power.



Syntax Tree
One way of parsing a (simple) sentence structure using NLP. Rather than a human manually reading thousands of customer responses, a computer can process the same amount in a fraction of the time.

How does text analytics work?

We've already discussed the importance of unstructured data in gaining a true understanding of your customers. But how do you efficiently draw actionable insights from something seemingly-unmanageable? This is where the power of text analytics comes in.

What is Natural Language Processing (NLP)?

NLP has actually been around for a long time and is suspected to have started in the 1950s with the Turing test. Since then, new technologies have emerged that aim to facilitate deeper and more comprehensive understanding of text data. NLP, specifically, is a computer algorithm that seeks to create an understanding of human language as it is used by people in everyday life.

These days, the two most-common approaches to NLP are rule-based human-driven systems or statistical machine learning systems. Both systems seek to create meaning out of the data through categorization - such as through tagging or bucketing of specific conversations - but use vastly different approaches.

Rule-based NLP is one of the older approaches, and is fairly simplistic. A good example of this would be a system that automatically tagged tickets based on certain pre-set words that a human being had selected and input.

Machine Learning NLP is a bit more advanced with a couple of different varieties, but if done right, promises significant efficiency gains over the traditional rule-based approach. We'll talk more about that later.

The move away from manual coding to rule-based NLP

Let's think of a pretty common issue: imagine that you work for a women's clothing retailer and want to categorize customer conversations based on the customers' response. These categories might include style, fit and shipping—manually tagging (often called **manual coding**) all of those conversations would take an incredibly long time.

Using a rule-based NLP approach to help apply categories to the data with the aid of a machine would speed up the process, but you would still need to have an analyst to create a set of categories that you would like to use and then manually create the rules to sort the text data into them. The analyst

would also be responsible for determining if and when those categories needed to be updated based on changing business needs or strategy. That would be called “updating the codebook”.

It’s still quite a bit of work, as you can see, but much less than if someone had to do all of it on their own without the assistance of machines. The problem with a rules-based approach like this is that it comes with inherent challenges and limitations, including human biases, time to results, and complexity of the process.

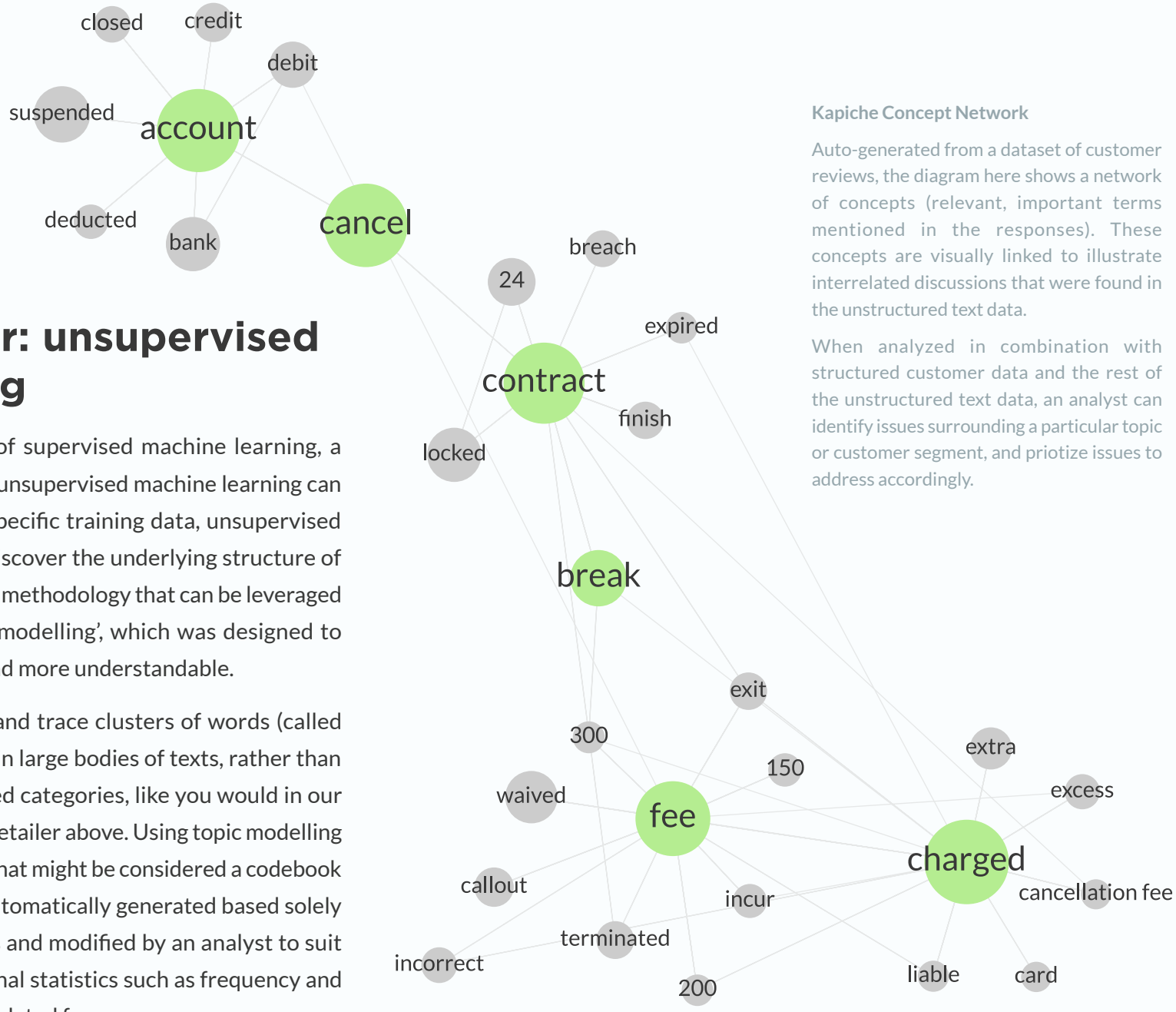
Enter supervised machine learning

Recently, new technologies have emerged in the machine learning domain that aim to facilitate deeper and more comprehensive understanding of text data with a fraction of the human effort. The amount of human effort required ultimately comes down to the specific approach taken.

The first and arguably still most popular approach (as of 2019) in this class of technologies is supervised machine learning. With this approach, you still need to provide the set of categories you want to apply. You also need to provide examples of where a human has applied these categories manually to existing data (this is called training data). You often need 10,000+ examples to train the machine learning algorithm. The more training data you provide, the more accurate the final result. The machine then learns how to apply these categories to unseen text automatically. Once you get this process right, you have eliminated a large chunk of the manual human work required in the traditional rule-based approach to NLP.

Unfortunately, it isn’t all good news. While on the surface it seems inconsequential to accumulate the required training data, it can actually prove to be quite a barrier. Additionally, while the training process can best be thought of as automatic, it takes time and often several retries to get right.

Then there is the issue of emergent or changing discourse. Just like in the rule-based example above, when new categories have to be added to capture a change in discourse, you need to re-train your model with more training data for that new category.



Kapiche Concept Network

Auto-generated from a dataset of customer reviews, the diagram here shows a network of concepts (relevant, important terms mentioned in the responses). These concepts are visually linked to illustrate interrelated discussions that were found in the unstructured text data.

When analyzed in combination with structured customer data and the rest of the unstructured text data, an analyst can identify issues surrounding a particular topic or customer segment, and priotize issues to address accordingly.

The next frontier: unsupervised machine learning

To overcome the shortcomings of supervised machine learning, a new but related technique called unsupervised machine learning can be used. Rather than requiring specific training data, unsupervised learning can instead be used to discover the underlying structure of unseen data. Specific to NLP, a key methodology that can be leveraged is 'topic modelling' or 'language modelling', which was designed to make text data more accessible and more understandable.

Topic modelling is a way to find and trace clusters of words (called 'topics' or 'themes' in shorthand) in large bodies of texts, rather than specifically looking for pre-defined categories, like you would in our example of the women's clothing retailer above. Using topic modelling technologies, a language model (what might be considered a codebook in manual coding terms) can be automatically generated based solely on the data uploaded for analysis and modified by an analyst to suit the business needs. All conventional statistics such as frequency and correlation are automatically calculated for you.

One of the other main benefits of emerging unsupervised text analytics software is the control it affords analysts in how they report insights out to the rest of the business.

A quality unsupervised text analytics system will show the analyst all the concepts and/or themes that are contained in their data rather than depending on the analyst to already have systemic knowledge. The analyst's job is to then decide how to best structure those themes for reporting purposes. Unsupervised machine learning can help you uncover what you didn't even realize you needed to look for.

In the case of our women's retailer example, this might be joining size and fit into one theme called "sizing." If the issues customers are talking about changes over time, the system will automatically detect and report these changes without any manual intervention from the analyst. This is infinitely better and more insightful than a manually maintained codebook of categories. Combining these features ensures organizations can track issues that are most important to their business, without missing important new issues as they emerge for their customers. It also means that the analyst can focus on making the system better as the AI's knowledge grows, rather than just trying to make the system effective in the first place.

Better still, depending on the unsupervised solution you decide on, little to no setup is required. The algorithm, once you get it going, will begin to pick up recurring themes in your data in minutes rather than days or weeks. This method also allows for a much higher degree of control over the language model that will then evolve over time with the business—your analyst won't constantly be tied to their computer to update the codebook any time your strategy changes and you want to track a new category.

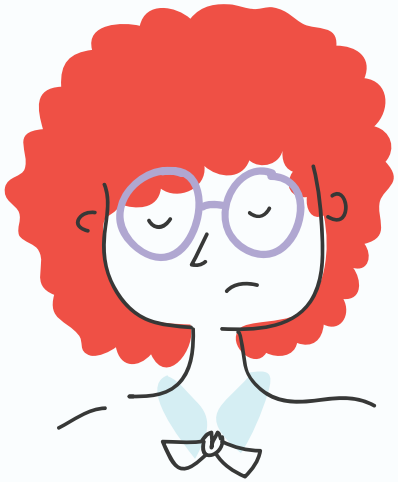
Whether the coding process is entirely manual or computer-assisted, the result is the same: you get frequency and co-occurrence statistics of categories which you can examine to better identify insights.

Strengthening your strategic decision-making using data

Text analytics are the key to getting the most out of any data you have collected, and allow you to deeply simplify the process of analysis. Unlike manual analysis, the process is repeatable and objective: different people feeding the same data into the system multiple times will all produce the same results. Not only does this provide solid ground for decision making on how to improve the customer experience, it also allows you to monitor changes in your data over time.

The ability to have an overarching, long-term view of your unstructured data is where organizations realize long term gains from machine learning and AI. By analyzing the trends in topics, both positive and negative, you can build a picture of how your customer experience team is performing and how strategic changes are affecting customer satisfaction. Better still, you can then use this data and correlate it with other key business metrics. Now that you've created this link to determine your ROI, future CX projects can be advocated for and developed based on hard data rather than gut feel.

Read more: Building a business case for CX in your organization



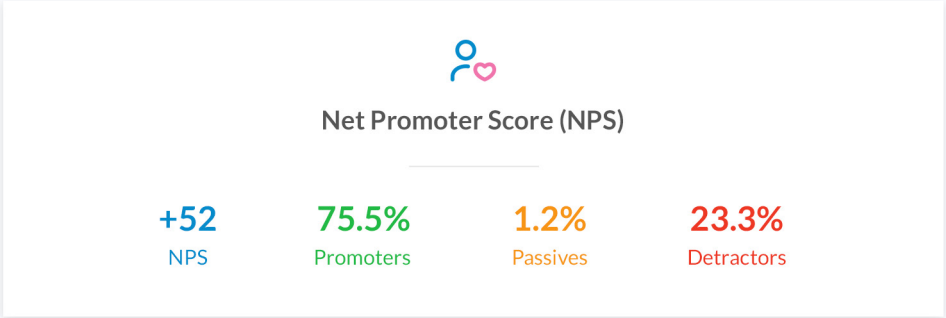
Adding some emotion with sentiment analysis

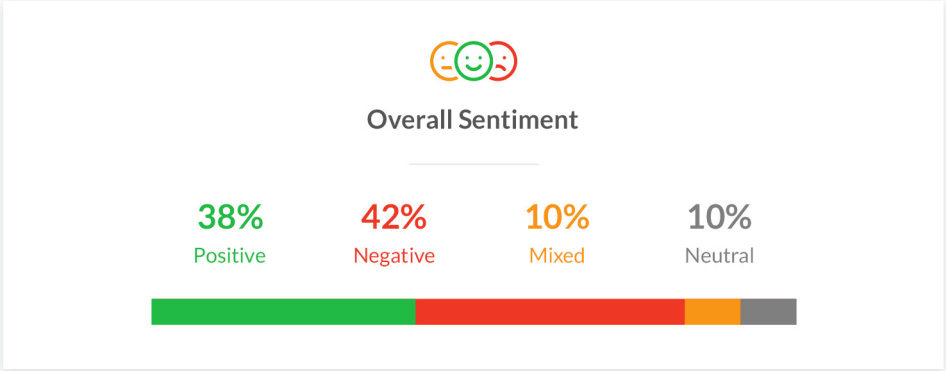
Sentiment is a great tool to assist with the analysis of text data and can give you a new dimension for seeing how your customers feel about your brand and product/service. It can help paint a strong picture of your business' customer experience and, when combined with text analytics, can be the difference between knowing that lots of people are talking about your new product and knowing that lots of people love (or hate) it. To get the most powerful results, though, it needs to be done properly. Here are a few things to keep in mind...

1. Build a framework around sentiment

Before using sentiment analysis, it's important to understand if sentiment analysis will add value to your existing data. For instance, if your existing surveys are filled with questions such as "what did you like best about your experience with us" or "what can we improve upon", then you can already see that the first will produce positive responses while the second will only list negatives—you don't need sentiment analysis to tell you that. However, if your data uses open-ended questions where you aren't leading customers to certain types of responses, then sentiment analysis can improve your understanding of the customer data. For example, if you asked "What would you tell a friend about us?" you would be in a prime position to see benefits from sentiment analysis.

It is also important to consider how sentiment analysis will be used as a metric. Will it be the primary driver for initiating CX change within the business or instead be used to enrich the insights from other, more concrete measures such as Net Promoter Scores (NPS) or CSAT? Treating sentiment as an enrichment to your existing data prevents you from being led astray by your sentiment model. For example, a highly satisfied customer might say 'I can't get service this good anywhere else!' which would be mislabeled as negative using the Google Natural Language API. Yikes!





Always try to combine sentiment analysis with a structured scoring metric such as NPS, to prevent tunnel-visioning on a potentially incorrect type of analysis.

2. The limitations of sentiment analysis

As far along as computers have come in with NLP, there are still limitations on what they can achieve. And it's not just their ability to comprehend written word and context that can derail sentiment analysis: one of the key things about sentiment is that it is highly subjective, not just across demographics but even across your own environment. For instance, if a customer is having a bad day, they are much more likely to give a strongly negative response to something that might have been neutral on any other day. That's why it's important to aggregate sentiment across as large a dataset as possible.

There are other aspects of human communication that computers still have trouble with as well, such as sarcasm and other ironic language. Without surrounding context, it is impossible for a computer to distinguish a proper "The service was fantastic" with a sarcastic "The service was fantastic". Given that, it's valuable to have another associated metric like NPS attached to a response. Pairing unstructured with structured data can assist by providing a comparison framework and backbone.

Lastly, we've previously touched on how your questions can affect the responses, but this is another area that can affect the potency of sentiment analysis. Consider the following response "Helpful staff that were there when I needed them". That answer would be interpreted differently were the question "How was your experience in the store?" or "How could your experience in the store be improved?". As you can see, the second question qualified the sentiment beforehand by using the formula of "How could XYZ be improved". To get the most out your sentiment analysis you should aim for neutral questions that make the respondent give unbiased feedback.

Examples of the pitfalls of sentiment analysis. Having associated structured metrics to use (such as ratings or NPS) help defend against poorly interpreting data based on sentiment analysis alone.

3. Include quality control

Before you dive in and start reporting with sentiment, it's helpful to confirm that your sentiment model is making sense of your data. Even the best sentiment models will make mistakes, and it's up to you and your analyst to make sure that these mistakes aren't going to impact the results you get.

While it's not in your best interest to read all of the responses (that would defeat the point of an automated tool), reading a sampling might help—particularly paying attention to any major themes and segments in your data.

Finding flaws and mislabelled examples doesn't invalidate your sentiment model or the utility of sentiment in general, but it does mean you need to pay close attention when interpreting and reporting the results. Your analyst might find particular segments where sentiment is inaccurate: perhaps you have a population of customers for whom English is not their first language, meaning just those segments need to be dealt with carefully. Or you might find that the sentiment output is less accurate across the board, meaning when you report sentiment metrics they need to be accompanied by a discussion of the precision of the results and how to interpret differences between segments.

Keep an eye out for:

Sentiment is unfortunately **far too often used without context or proper understanding**. Simply looking at sentiment results and assuming "lots of positive = good" is an incomplete and sometimes harmful conclusion. To get the most accuracy out of your data when incorporating sentiment analysis, remember these three points:

1. Make sure the verbatims are responding to neutral, non-leading questions.
2. Combine sentiment results with other structured metrics such as your NPS, CSAT or other score systems.
3. Read a sampling of your data to make sure there's no obvious issues with the sentiment labelling on your particular dataset.

PART TWO

Using text analytics to achieve organizational goals

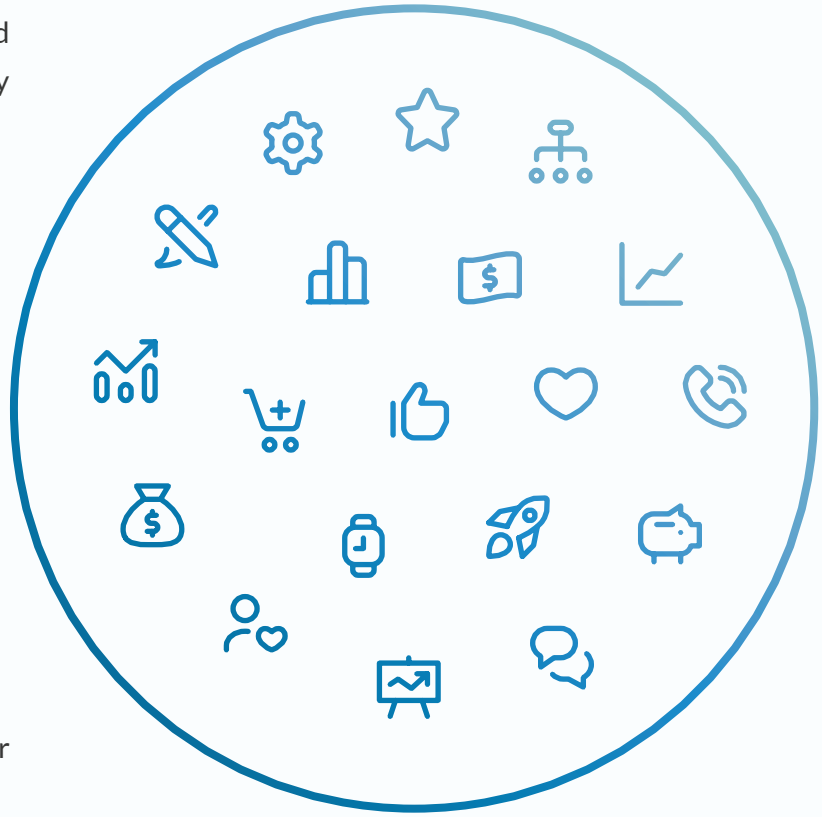
The human touch.

With today's text analytics tools, analyzing data has never been easier. However, to ensure changes are implemented to improve customer experience, the information needs to be disseminated to the relevant people in the organization.

A good customer analyst will have strong relationships with senior executives and the ability to take the business on a journey using the data to illustrate the importance of strategic CX.

What are you trying to achieve?

Organizations typically first encounter text analytics as a part of an end-to-end Customer Experience (CX) platform. CX platforms incorporate basic text analytics software as a part of a much broader solution that covers customer engagement, surveying and communication—a great example of this would be a tagging system that automatically tags conversations based on certain words or phrases.



Unfortunately, though, these end-to-end CX solutions come with one major disadvantage: organizations are confined to using the text analytics component purely for analyzing customer experience data. So, if you wanted to apply text analytics to data from other parts of the business (for example, employee experience, L&D and operations), it is impossible to do so. In that case, you would still need to use a specialized text analytics solution separate to your CX platform.

Before you purchase any text analytics solutions, take the time to make sure it will achieve the goals your organization has set. There are many different platforms and software packages available that all have their pros and cons. Ending up with the wrong solution could mean additional money and time delays, preventing you from taking fast action to improve your organization's customer experience. Here are a few other things to consider as you analyze your options...

Level of insight & accuracy

Open-ended customer feedback is one of the most valuable data-points your organization has. Feedback on how your customers feel about you provides valuable insight into what is working and what isn't. When analyzed properly, this feedback gives you the exact information you need to improve CX. Doing so will give you the even better bonuses of subsequently building loyalty, increasing average customer LTV and growing referrals to new areas of your business.

That being said, many analytics solutions currently on the market reduce this valuable data to a set of codes and analyze it as if it were any other type of structured data, causing organizations to miss valuable information and the opportunity to empathize with their customers.

The problem here is search vs discovery.

Constrained by historical methods, analysts are searching for things that they already know they need to look for and are missing out on the emergent issues that their customers are providing feedback on.

This puts organizations on the back foot, unable to accurately see and fix the problems customers are facing until they become a much bigger issue. This puts organizations on the back foot, unable to accurately see and fix the problems customers are facing until they become a much bigger issue. However, recent advancements in text analytics solutions remove this issue by allowing analysts to discover which topics or issues customers are talking about and to identify emerging trends in their data as they arise for customers.

Kapiche improves on this further by identifying influential concepts and then providing the much-needed context around these concepts to paint a complete picture of the customer's feedback. Kapiche does this via an algorithm that builds a unique language model, without input from the analyst. Without the need for human input, assumptions about what is important to customers and inherent human bias are removed, ensuring a much more accurate picture of the data.

Time to results

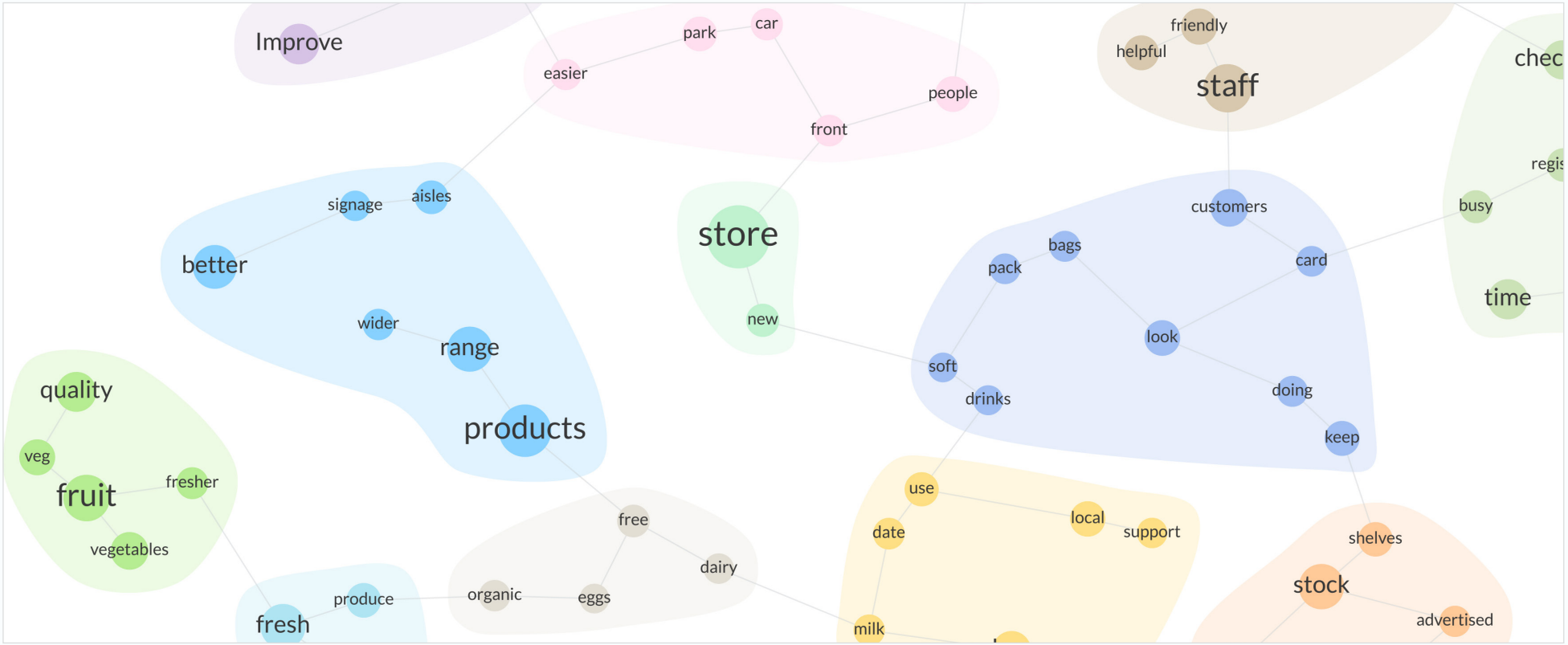
Before considering a text analytics solution, most organizations will try the manual approach—manually reading each verbatim and hand categorizing it. Sometimes this is because of cost, or perhaps lack of understanding, but often it's because companies don't know there's something better out there. Whether it's done internally or outsourced to a 3rd party, this approach has a number of significant limitations. Firstly, the time to results is often measured in days or weeks, rather than minutes. Similarly, a solution that requires a human to manually tag everything isn't scalable for the long-term success of your team.

Worse still, some solutions try to replicate this process but replace the manual reading with machine reading. Often, these systems still require an arduous setup process (from days to months) that has to be repeated for

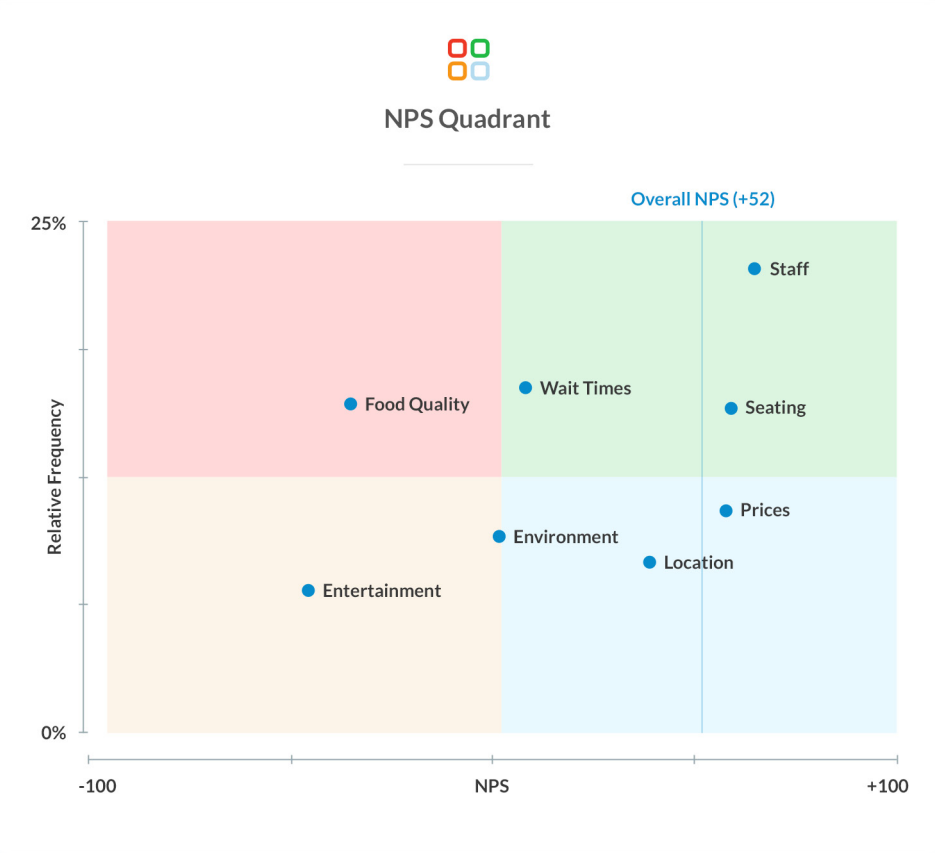
each different type of data analyzed, significantly impacting scalability and time to result. It almost defeats the purpose of using a technology solution in the first place.

Luckily, recent advances in AI-powered text analytics address this problem using unsupervised machine learning. These types of solutions dramatically decrease the time to results. Organisations are able to analyze 1,000s of verbatim comments in seconds with zero setup, as well as being able to go back and run analyses on historical datasets to track progress over time.

No matter which solution you choose, be sure to consider how long will it take to start seeing results. Consider also how much effort it will take and how long it will be before you use this solution on a different data domain like internal employee surveys rather than customer surveys.



Above: Part of a topic model from 17,500 survey responses. Each node is interactive and directs to more information along with the matching responses. The visualiation was generated in approximately 30-45 seconds – many times faster than a human manually reading and categorizing.



A quadrant chart for a customer feedback survey showing numerous topics of interest with respect to the overall NPS and frequency. We can see from the above chart that "Food Quality" is an area of concern, located far below the overall NPS and with relatively high frequency.

Easily consumed output

Text analytics software is often used by several people within your company—both directly and indirectly. The first and foremost person using it is the analyst. This person deep dives into the customer feedback data, works to pull out and understand key themes from in-depth analysis and subsequently creates visualizations to tell a story for the rest of the business.

However, having an analyst identify actionable insights from the text analytics software is, by itself, insufficient to inspire change. Analysts must be able to share insights in a way that is easy for other teams to understand, otherwise they'll never be able to get buy in and make the necessary changes to improve customer satisfaction.

For example, executives want to see a high-level overview of key issues facing customers, factors that are positively and negatively impacting customer opinion or NPS and potential changes that need to be made in order to improve customer satisfaction. As such, it is important that these insights can be shared—including to those who aren't analytically minded—without forcing them to dig through the data themselves.

The added benefit here is that the ability to share insights across the business builds additional cross-business buy-in—one of the biggest contributing factors to whether an organization's customer experience program will succeed.

Trial using your own data

The above requirements are vital to choosing the right solution for your business, but how can you know what each solution is all about without testing it first? Considering the sizable investment that some solutions require, having access to trial the tool prior to purchasing ensures the software is fit for your needs and can deliver ROI.

Current text analytics solutions almost always have a demo available. Understandably so: they typically require large amounts of customization or set-up. A demo with a sample dataset will allow you to visualize how the program works and how actionable insights are discovered without having to wait for setup to take place to suit your requirements.

Some solutions will provide you with generic datasets and analyses predefined for a particular industry. Identifying the actionable insights specific to an industry dataset can be challenging when you don't understand the industry, so look for opportunities to use your own data. Kapiche choses to provide a full access trial that allows you to use your own data. Working your way around the tool is the best way to fully understand how it will (or won't!) aid your business, and see what it might look like to have it as part of your process. Having access to a trial like this should allow you to validate what you already know or suspected about your data, while also presenting you with new, actionable insights that you weren't previously aware of.

It is important to note that the most critical requirement, beyond all of the above, is having clarity around your targets and objectives of using text analytics. Knowing the amount and type of data you will analyze, how much you have to spend on the solution and how much time you want to put into setup and maintenance (that could be spent digging deeper into the 'why' behind your customer feedback!) will ultimately impact which solution you decide upon.

How to choose the right solution

We see many organizations struggle to choose the appropriate system for their purposes. Rushing into the purchase of a software product without understanding what is required to make that system work can leave you with a system that doesn't fit your business and no budget left to actually improve anything. For example, if you're considering a standalone text analytics solution, consider its ability to integrate with surveying tools such as Qualtrics, GetFeedback or Delighted so that survey data can be easily analyzed without the fuss of exporting and importing.

To help you figure out the best solution for your business, we've highlighted five questions to ask to help you navigate the procurement process.

Do you have the right data?

If your organization is collecting limited data, or the wrong type of data to improve your CX, no software is going to help. Before you start looking at potential solutions, consider the data you currently have available - or the data you plan on collecting. Is there, or will there be, sufficient data to tell a complete story about how your customer's experience your product or service? If not, output from your text analytics software will be limited and your organization's ability to make a difference to your CX program is left to gut feel rather than hard data.

To make any text analytics solution work for you, survey questions should be open ended to prompt the respondent to respond in detail, rather than a two or three-word answers. Questions should have a neutral tone: asking what a customer likes best about a service or what an employee thinks the organization should improve colors the responses either positively or negatively. Instead, leave it as open-ended as possible so the respondent can enter whatever is on their mind. So, instead of asking "How much did you love our support?" instead ask "How did you feel about our support?"

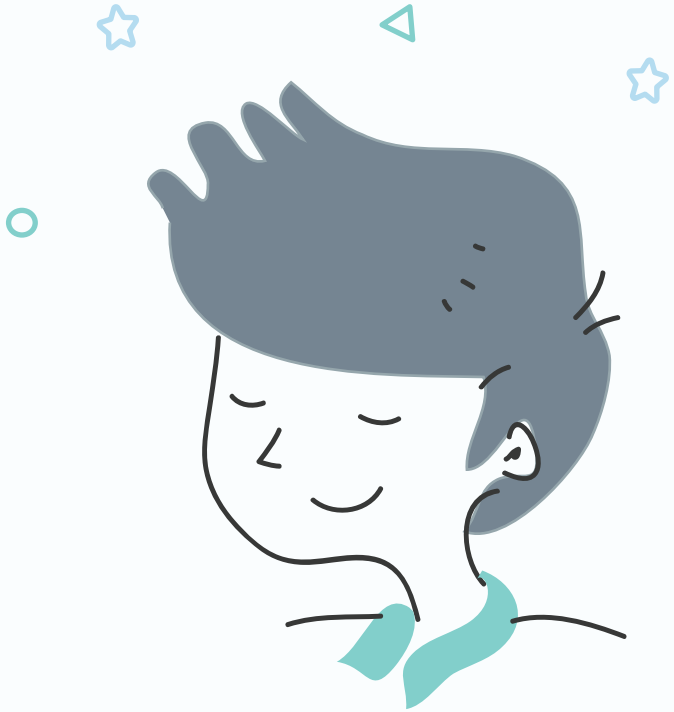
Something else to keep in mind: if you are receiving less than 100 survey responses per month, text analytics may not be the right solution for your business.

When to switch to text analytics software?

Analysing small amounts of data is useful, but text analytics generally deals with larger, more unwieldy amounts of data. There's no point in purchasing a whole text analytics solution for data that could simply be manually read in compareable amounts of time.

Here's some quick guidelines on when to consider looking into a text analytics solution for your organization:

- You are collecting over 150 survey responses a month
- The text responses gathered are descriptive and contain more than just a few words
- You are gathering structured customer data or scoring metric to use alongside the unstructured text data



Do you have the right people?

Text analytics software is designed for just that: analysis. It takes the effort out of combing through reams of data, but it still requires an analytically-minded person to make sense of what the data is saying in order to discover the actionable insights.

Typically, the right person for the job is a Customer Insights Analyst or someone in a similar position, with exceptional data skills and the ability to translate these insights into actionable insights for the relevant parts of the business.

To see ROI on your text analytics solution, an analyst should be optimizing, analyzing and querying data to find actionable insights and influence customer experience strategies across the business. Even if you are using an elaborate machine learning service, there should still be a human touch—preferably one who has done something similar before.

Do you have business buy-in?

Customer experience doesn't start or stop with the CX department alone. Without commitment from all departments to make the necessary changes to improve customer satisfaction, a CX program will fail before it has a chance to really make a difference.

Acting on the insights that you receive is where the ROI will come from. A recent study by Harvard Business Review found that 58% of enterprises are seeing a significant increase in customer retention and loyalty as a result of using customer analytics. But this wouldn't be happening without buy-in from all areas of the business. An analyst can pull out the most insightful, actionable insights from the text analytics tool, but without the right people willing to make the necessary changes to systems and processes, nothing will get better for your customers.

It is also important to note that, just like CX, text analytics should not be considered by your business as an ongoing drain of time and resources. On the contrary: when implemented and used as intended, a solution like text analytics software becomes part of your business and should provide ROI in the form of actionable insights almost immediately after implementation.

Will the program deliver valuable wins, quickly?

As more and more organizations realize the need to move away from product-centric business models and put the customer at the center of their brand, they don't see the impact as quickly as they would like to. Without every person committing their team-wide efforts, they struggle to find the time to create an amazing customer experience. But customer experience doesn't have to be hard, or slow to implement!

To kick off a CX, EX or VoC program in your organization, focus on quick wins, as these will be instrumental in demonstrating the value of your text analytics investment and should help you to start seeing a positive shift in your CSAT, CES or NPS score. They'll also seem less high stakes to members of other teams that you need to ask for help. Ideally, a text analytics solution can be used instantly with the data you already have, without the need for heavy customization or weeks of set-up and training.

Start by focusing on the low-hanging fruit that can be easily addressed. These are the areas most frequently identified by your customers or employees as needing to be improved. Once the quick wins have been attended to, the organization's focus can be turned to more strategic approaches to improving CX. These could be things like combining insights with your financial and transactional data to give you a detailed look into what your customers are 'doing' versus what they are 'saying' and where the disconnect lies. Read more about this here.

> "Data is a key transformation tool for all departments. It has become a core way of working. It drives a learning & innovative culture with clarity on contribution to the business (therefore allowing additional budget request when there is a business case), less prejudiced decision making and systematic, continuous improvement.
>
> — *The Customer Experience is Written in Data, 2017*

Will the solution suit your organization's needs?

We see many organizations default to thinking that a complete CX software platform—one that covers every aspect of the CX process from CRM to survey and analysis—is the right option. The downside is that, as with any generalized solution, you are trading specificity with simplicity, often at the cost of an enterprise level price.

Often, text analytics forms one small part of these larger CX software platforms. For very large organizations, it may make sense to implement a system that does everything. But for organizations that already have CRM and surveying tools covered, or are looking to have the biggest impact on their CX without paying out tons of money, a dedicated text analytics solution is a powerful way for an analyst to dig deeper into their unstructured customer feedback data.

The added benefit of a dedicated text analytics solution is that there are no limitations to the type of unstructured data that can be analyzed. Consider how other parts of the business, aside from CX & EX teams (tech, product, support & safety) could utilize text analytics to improve their departments. There are so many additional applications outside of CX!

GENERALIZED CX PLATFORM	SPECIALIZED TEXT ANALYTICS
All-in-one platforms trade specificity with simplicity at the cost of enterprise level price	Specialty text analytics focus on one area and may require supplementary products.
The text analytics component is often underdeveloped, making up a small part of the package	Dedicated solutions will be more robust, accurate, and fully-featured in their field
The text analytics component is limited to use in CX teams alone, unable to be utilized by other areas of the business	Some specialized text analytics solutions allow organizations to utilize their service across other parts of the business

Whatever the solution, make sure it suits your organization's needs

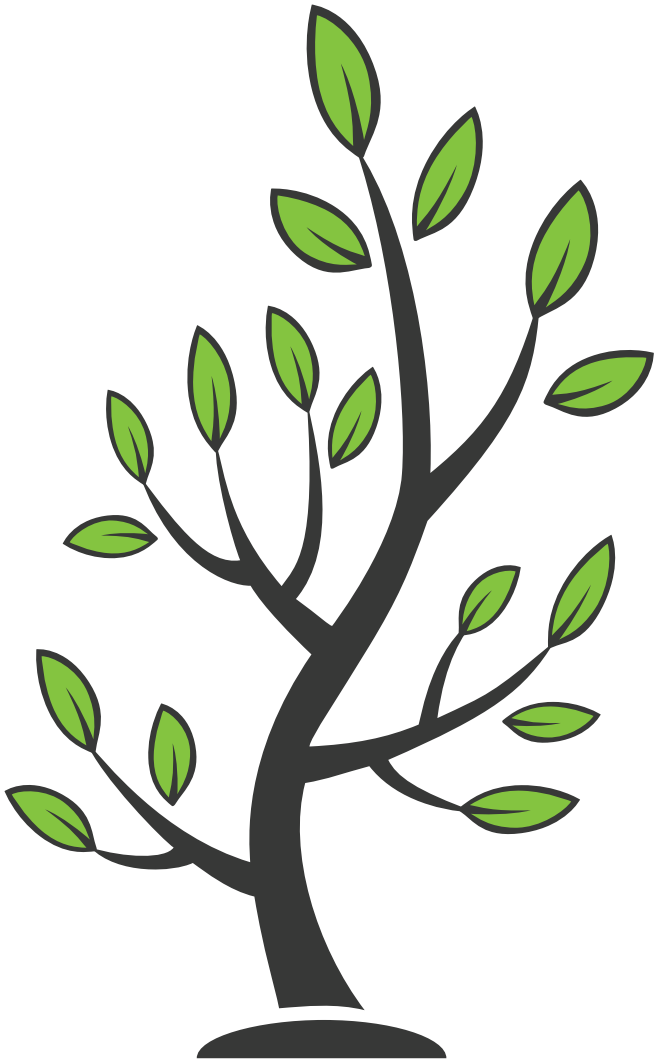
The purpose of text analytics for CX is to deliver valuable insights quickly, without huge investments of time and manual labor. The process should not be a burden—it should be easy both to implement and to use.

Whether your organization is considering a complete CX platform or a dedicated text analytics solution, it is important that everyone is clear on what you hope to achieve by implementing it. This clarity will ensure the chosen solution is the right fit and will help to achieve your goals.

Ensure that the right people are onboard and dedicated to making the changes necessary to improve customer satisfaction—with the right people in the right departments driving a move towards customer-centricity, your organization is likely to greatly improve customer satisfaction, customer loyalty and ongoing growth.

PART THREE

The Kapiche Approach



The Kapiche approach

Kapiche helps organizations increase their revenue and improve customer satisfaction by enabling them to quickly identify and address those customer issues that have the biggest impact on revenue.

A great example of this is one of our customers who, before Kapiche, was only able to analyze two responses per minute out of their 5,000 monthly survey responses—on a good day. Since using Kapiche, this customer is now able to analyze all 5,000+ responses in minutes, as well as being able to go back and run historical datasets to track progress over time.

Another customer has been able to decrease their analysis costs per survey from $3.00 down to less than $0.10.

Check out how Schindler manage their unstructured data using Kapiche

Our focus is on unsupervised, statistical text analysis. As detailed in Part 1, this mode of text analytics removes the need for resource intensive machine learning by allowing users to upload thousands of survey responses and analyze them in seconds. It also allows you to identify and act on emerging issues as they arise.

Kapiche specializes in taking the data from open-ended unstructured questions along with data from traditional close-ended structured question to give deep actionable insights, which we believe come from this cross-section of structured and unstructured data. We believe this gives organizations a high degree of visibility that allows them to make strategic decisions to increase customer loyalty, growth and revenue.

Understand customer opinions, improve customer satisfaction

Quickly understand what your customers think about your company – in their own words – from your NPS®, CSAT, customer survey and product review data.

www.kapiche.com