

TEXT ANALYTICS

DERIVING ACTIONABLE INSIGHTS FROM UNSTRUCTURED TEXT

SUIT THE ACTION TO THE WORD, THE WORD TO THE ACTION
"Hamlet Act 3, Scene2"



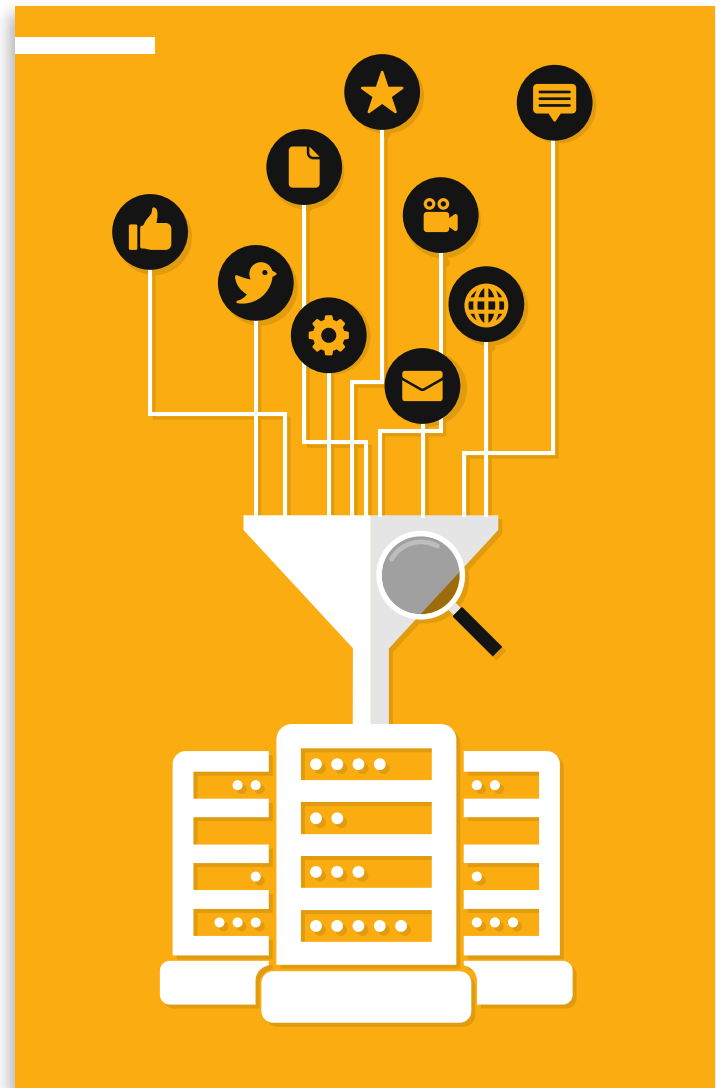
WHY INVEST IN TEXT MINING?

Unstructured data is the most prevalent form of information on the planet. Vast amounts of new information with significant economic value are generated everyday through different business activities. 80% of this information lies in the text form across reports, articles, survey data, call centre notes, business mails, social media sites, local business websites, chats, voice transcripts etc. However, most of this information is hardly even used to obtain better insights into customer behaviour or aid strategic decision making. Most companies do not have the right toolset to organize this form of unstructured textual data for future use.

Suitable extraction and analysis of this data can feed into the master customer database of a business, making its bank of consumer insights richer, insightful and more robust. Mining unstructured data however, is a daunting task. Unlike the methodical structures in which information is arranged in surveys and retail POS data, leading to easy processing and reporting; unstructured data is fluid and open ended.

Text must therefore be accessed, copied, analysed, annotated and related to existing information to make it intelligible. However, more and more businesses are investing in the analysis of unstructured text data. Why is this happening?

Text mining helps an organisation bridge the last mile gap between the business and the customer. It helps the business “listen in” to customer conversations and design customised engagements for them. It builds the brand image as proactive and agile. It helps the business create personalized discounts for the customer, stay informed about differential pricing which customers expect from online-offline trade or competitors, get feedback on newly launched products and services, identify and service immediate business opportunities.



The benefits of this strategy are clear. By “listening to the consumer”, by creating personalized product and service offerings, by discontinuing or improving product/service lines which receive negative attention, by being agile with price points, retailers increase their chances of making a sale, even making a greater profit on specific products. On the other hand, customers receive deals on the products they are interested in. Studies also suggest that customers feel special when their brand “listens and reacts to their suggestions”, which encourages brand loyalty. Text mining practices improve the entire customer experience from engagement to post-purchase.

WHAT DOES

AQ BRING TO THE TABLE?

AQ has helped global clients across industries answer different business questions with our text mining abilities. The highlights of some of our key projects are captured below:-

THE CLIENT CHALLENGE

A popular US manufacturer of speciality coffee and coffeemakers wanted to understand why an upgraded variant of its already successful coffeemaker was receiving a lot of negative response from consumers.

THE AQ SOLUTION

We mapped themes relevant to the coffee industry to each review and calculated the sentiment expressed by consumers before and after the launch of the upgraded variant. This helped us identify and answer two important questions; had the overall net sentiment about the coffeemaker really declined and what were the reasons behind this decline.

Our analysis showed us that the net sentiment about the coffeemaker had truly declined. The reason behind this decline was due to the consumer concern that the newly launched coffeemaker variant was limiting consumers to using only company manufactured coffee pods. Our analysis helped the company on their path of course correction.



THE CLIENT CHALLENGE

A leading cruise company wanted to understand which aspects of its services drove its ratings

THE AQ SOLUTION

To help answer this query we broke down the company's reviews into themes and ratings, dividing them into four buckets. The themes were;



Hygiene

services, which if neglected would pull the ratings down.



Must Have

services, which were highly impactful and influenced low/high ratings.



Low Priority

services, which had the least impact on overall ratings.



Delighter

services, which if rated higher would increase the overall rating

We then presented the client with a focus list of service areas. Our analysis also revealed that our client needed to improve on themes like ‘Dining Experience’, ‘Cabin Area’, ‘Fitness and Recreational activities’ and invest in markets like Europe and Alaska. Our analysis helped the company invest in initiatives which improved its service line.

THE CLIENT CHALLENGE

One of the largest automobile manufacturers in the world wanted to understand the expectations which customers had from their dealers and gauge their levels of satisfaction. They also wanted to identify dealer attributes that drove consumer ratings.

THE AQ SOLUTION

We classified the client’s online reviews into “sales” and “service” buckets and analysed them separately. Analysis of the sales reviews gave us the insight that “ease of process” and “monetary aspects of a deal” were the two main keys to customer satisfaction. In the service reviews category, “dealership facilities” and “car preparation”, though less talked about, were influential in improving ratings. Our analysis helped the company optimise its sales and dealer operations.



THE CLIENT CHALLENGE

A US pizza brand wanted to understand the main drivers behind consumer ratings on social media and identify areas of strength and opportunity for the brand.

THE AQ SOLUTION

Our process involved mapping themes from reviews, followed by the calculation of sentiment scores for each theme. The overall insights were as follows. ‘Overall Service’, ‘Staff Behaviour’ and ‘Pizza Taste’ were the most influential driver of ratings. The pricing across the chain was perceived to be ‘Value for Money’, and that was identified as its chief strength. However, consumers were not happy with aspects like ‘Staff Responsiveness’ and ‘Staff Behaviour’. Our analysis helped the company understand that it needed to provide a better service experience for customers and invest in training its staff to be more customer-friendly.



THE CLIENT CHALLENGE

One of the largest chocolate manufacturers in the world wanted to analyze their customer reviews to understand the trend of opinions and sentiment towards their product lines.

THE AQ SOLUTION

On analysis, the themes of 'Meltability' and 'Improper Packaging' arose as two of the key drivers of customer sentiment. It was also found that customer sentiment varied with changing seasons and negative reviews about the brand increased during the summer. Our analysis helped the brand alter their packaging and delivery methods to prevent degrading of products in transit.



WHAT'S OUR RECIPE FOR SUCCESS?

AQ'S TEXT ANALYTICS SOLUTION

Extracting insights from unstructured data is a complicated process. In the table below, we have tried to illustrate the common challenges which emerge in this process and the ways in which we solution these problems. AQ's text analytics solution is robust, flexible and can be adapted across a wide variety of industries and business scenarios.

CHALLENGE



Data Quality

Unstructured data frequently contains repeated punctuations, errors and short forms. Consumers often mention multiple unrelated ideas and sentiments together and at times (especially on social media) write things that have no relation to the business or organization in question.

THE AQ SOLUTION



Data Clean-up & Handling

We clean the data using our in-house, heuristics based text cleaning scripts. This activity cleans unnecessary elements from the text that don't capture any information for eg; repeated punctuations, extra spaces, symbols etc. Industry specific short forms and uncommon words are corrected by adding new rules to the existing list. We also use other meta data captured along with the text data to classify reviews into groups so that unrelated reviews can be identified.

CHALLENGE



Making Sense of Natural Language

Unstructured data makes it difficult to identify topics/themes that are

| **GRANULAR**

| **RELEVANT**

| **ACTIONABLE**

It also makes it difficult to fit customer comments into their proper domain context.

THE AQ SOLUTION



Tailoring themes/mapping topics for the industry

Once the data is cleaned, a sentence token annotator from open source NLP packages is used to break down the reviews/comments into manageable sentences. A POS tagger then assigns a part of speech (Noun, Verb, Adjective etc.) to each word present in these sentences.

Then, AQ's proprietary algorithm parses these words and extracts useful phrases from each of these sentences. This forms an important base for the theme mapping stage that follows. Themes and keywords are selected in consultation with our clients according to their business context and requirements. A training set is prepared and the model is trained on a sample dataset to recognize patterns – which sets the framework for automatic coding.

Keywords and themes are then progressively tested on larger data samples and perfected over iterations till the desirable level of accuracy is reached.

This method of mapping themes at a phrase level helps us to achieve the required granularity for calculating sentiments and generating insights later.

CHALLENGE



Identifying satisfaction at the context level

A star rating does not always tell the whole story. It is important to segregate themes which are driving satisfaction / pain-points. For example, social media comments

THE AQ SOLUTION



Sentiment Analysis

Sentiments are derived using context clusters instead of simple, singular words. They help us understand and identify the semantic linkages among word clusters and

CHALLENGE

almost always show a positive skew but there might be underlying issues that a significant number of customers would appreciate being noticed or fixed.

THE AQ SOLUTION

provides weightages for words that are neutral, negations, amplifiers & de-amplifiers.

Sentiments are captured on a continuous scale (+1 to -1) from strongly positive to strongly negative. Our algorithm also considers generic and industry specific usage of words and phrases.

The granularity of sentiments ensures that we can pin-point any major concerns or issues that are expressed in comments.

CHALLENGE



Separating the “Noise”

Reviews and transcripts cover a wide variety of subjects, themes, opinions and satisfaction-dissatisfaction feedback from consumers. Not all this content is relevant, actionable or has any serious influence on the complete set of target customers for a business. Therefore, it is necessary for a business to separate the “noise” from the feedback and isolate the real areas of concern or improvement.

THE AQ SOLUTION



Actionable Insights

Our statistical modelling techniques measure the influence that themes and topics have on overall satisfaction scores. This, combined with the analysis of themes and granular sentiment scores makes our modelling technique unique and impactful. We can quickly identify, isolate and present not just problems but insightful solutions which the company can then implement. We can provide insights on key areas of improvement, identify areas of customer satisfaction, map consumer alignment across different regions, demographics etc.

WHAT ARE THE OTHER OFFERINGS ON THE TABLE?

AN ASSESSMENT OF THE AVAILABLE MARKET OFFERINGS

There are many players in the market who are trying to address the need to analyse unstructured textual data with packaged and preconfigured text analytics offerings. A “Forrester” study has evaluated the major players through the three lens mentioned below.



STRENGTH OF THE OFFERING

This metric evaluates the functional and technical capabilities of different solutions such as text mining based on linguistic and/or statistical rules, data enrichment with business-domain- knowledge for improved accuracy, GUI, architecture, scalability etc.



STRATEGIC PRIORITIES

This metric evaluates the players on their vision & strategy, intellectual property and partnership ecosystem

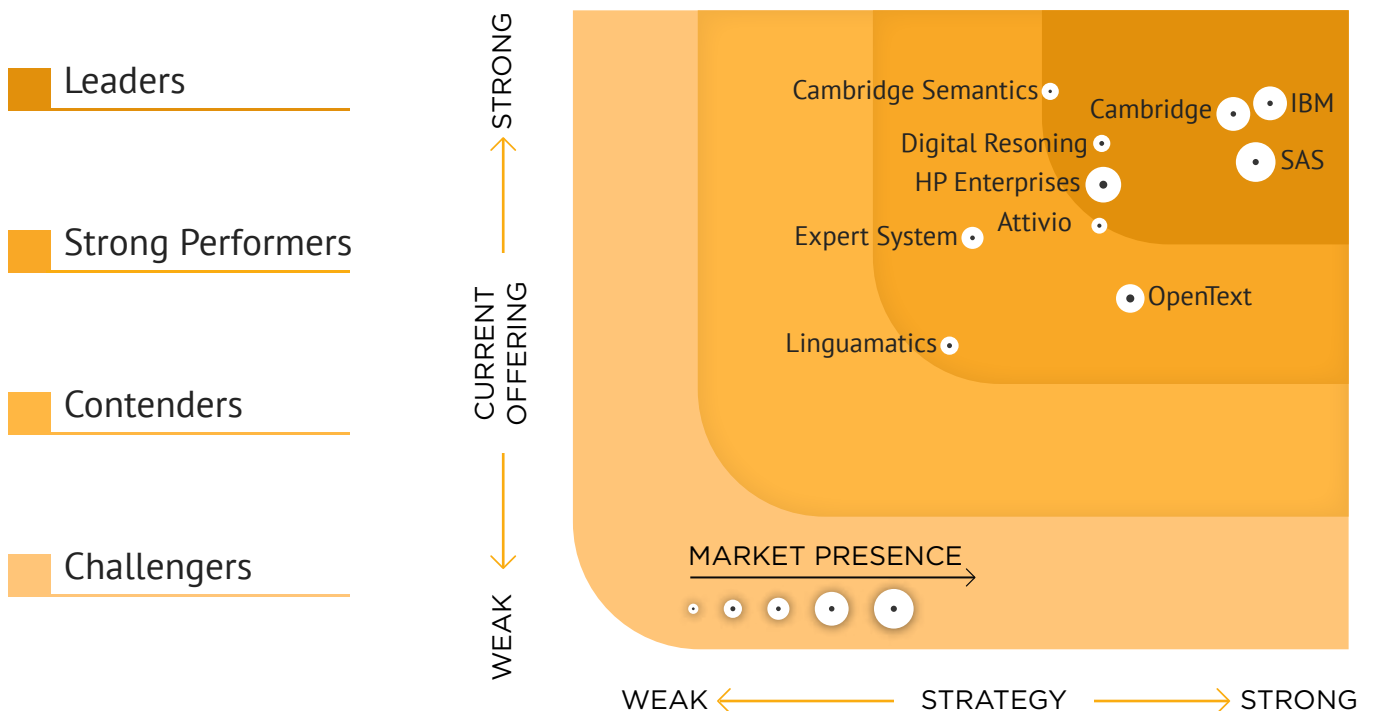


PRESENCE & REVENUE

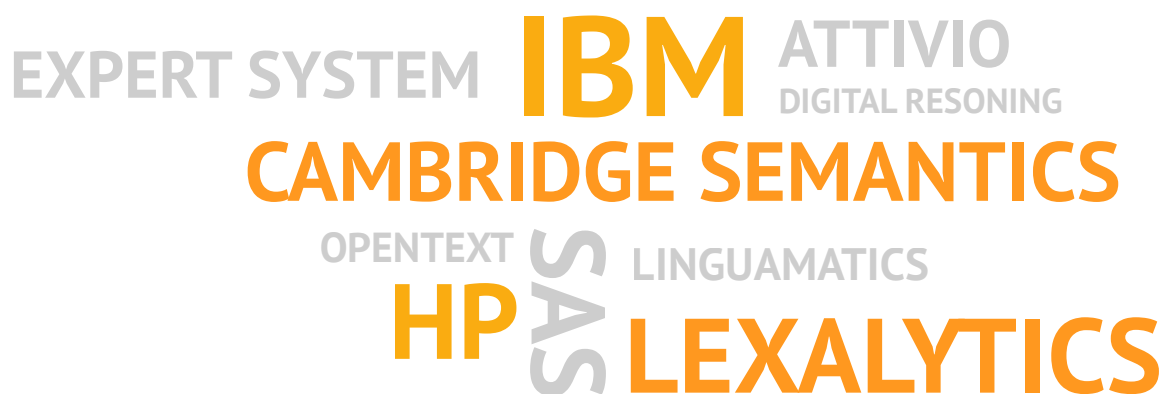
This metric evaluates the players on their installed client base, revenues and global presence.

The chart below showcases how a few of the big players in the market stack up against each other based on the strength of their offerings, strategic priorities and presence.

THE FORRESTER WAVE™ BIG DATA TEXT ANALYTICS PLATFORMS



SOLUTION HIGHLIGHTS OF THE KEY PLAYERS



IBM's Watson Explorer has preconfigured capabilities which go beyond linguistic and statistical rules and deliver out of the box value for clients.

Clarabridge has a strong text analytics platform, with its differentiator being its VoC and VoE applications.

HP's IDOL is a unified platform for text analytics which supports out of the box connectivity, taking care of data relocation risks and unnecessary storage costs.

Lexalytics engages a number of sophisticated text analytics systems that aim to answer questions around what people are saying and the associated sentiments.

Most existing text analytics platforms seem to be leaning towards higher automation and an attempt to eliminate dependency on human resources. However, a thorough analysis of unstructured data requires domain insight. A text analysis study requires prior knowledge of what the text is about and a great deal of upfront work to build relevant dictionaries or ontologies. The biggest risk in automating the analysis is that the process of extraction is biased towards previously chosen words and phrases and hence prevents specific granular insights from being revealed.

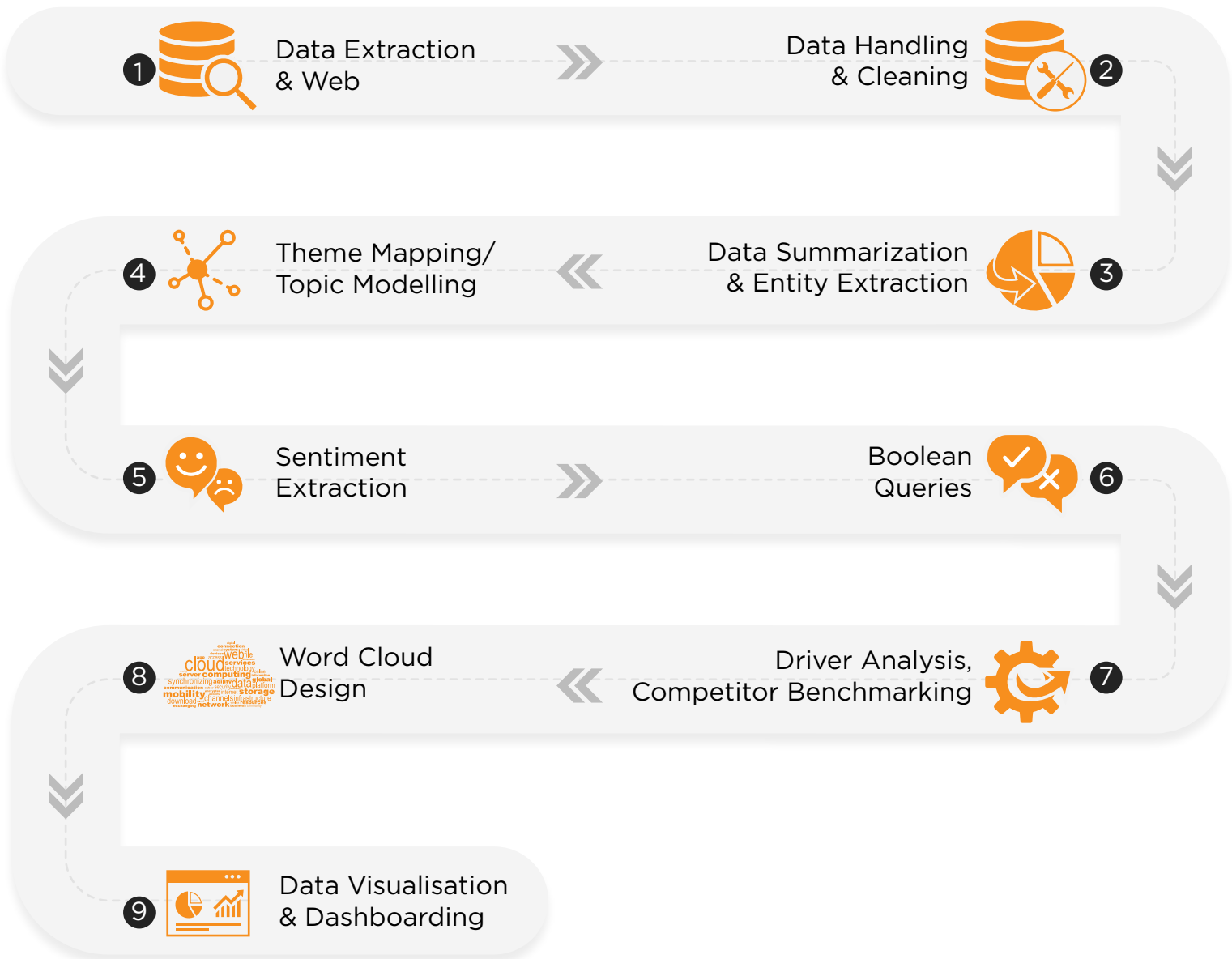
Many solutions claim to handle text. The truth is that they only “handle” it by storing and displaying it in full, but lack a true understanding of what’s inside: natural language. Search-based solutions typically cling to statistics about previous searches by other users and therefore are inherently biased with the ideas and reasoning of previous users. Some expert systems combine a machine interpretable knowledge base with a reasoning engine to draw conclusions from the text, but they are limited by the scope of that knowledge base as well, since that base requires a large budget for building and maintenance.

WHY IS OUR SOLUTION MORE HOLISTIC?

THE AQ DIFFERENTIATOR

We have done a large body of work in text analytics and are very proficient with the entire process from “data extraction” to the “presentation of insights”.

JOURNEY MAP OF THE AQ PROCESS :



Our solution offers true insight into text by starting from the language itself. Its unique approach translates free text into quantifiable entities with rich contextual annotations. All of this is solely based on the properties of the natural language in which the text was written and is therefore strictly unbiased. Our process enables us to identify, isolate and present not just problems but suggest insights which can help companies plan better.

THE TABLE BELOW SUMMARIZES A FEW KEY DIFFERENCES BETWEEN THE TWO APPROACHES:

	AQ'S SOLUTION	OTHER PLATFORMS
Services Provided	Text Mining, Reporting, Analysis and Visualization	Predominantly Text Mining and Reporting
Accuracy	High (>85% accuracy)	Average in most cases
Supervised and/or unsupervised techniques	Supervised (Customized themes and other requests can also be incorporated)	Unsupervised (Standardized results restrict deep analysis)
Drivers Analysis	Determines the impact of different aspects of the business on consumer behaviour	Service not provided
Actionable Insights and Recommendations	AQ's solution can provide granular and actionable insights	Service not provided
Visualization	Completely customizable. Visually rich and more intuitive than any non-customized market solution. Can be accessed from anywhere using a web browser and internet connection	Standard charts and tables
Implementation Time/Cost	Low to Medium	Low
Licensing Fee	Nil	High

WHAT'S THE PROOF OF THE PUDDING?

HOW DO WE IMPACT BUSINESSES?

WE IMPACT BUSINESSES



WITH OUR ABILITY TO MARRY OTHER DATASETS WITH TEXT ANALYTICS TO CREATE A HOLISTIC STORY AND ANSWER RELEVANT BUSINESS QUESTIONS

Our text analytics solution is highly flexible and can be integrated with different types of data sets. Analysis findings can be clubbed with different types of datasets like POS transactional information, store details and performance, brand level KPIs, financial datasets etc to answer questions like:

What is the relationship between customer sentiment and spends on products or types of products purchased?

Is there any relationship between the performance of my marketing/promo campaigns and consumer sentiment?

What are the key drivers of customer satisfaction? How can I track performance on these key drivers and analyse its impact on my top line?

What are my best and worst performing stores? What are the underlying reasons behind their performance?

How do I improve the social health of my brand, which key factors do I address first?



BY SOLVING SPECIFIC PAIN POINTS WITH STANDALONE ANALYSIS

We can help clients with targeted standalone analysis and solve specific pain points like;

Evaluating brand performance by generating insights from an analysis of verbatim comments

Identification of customer concern areas

Tracking competition

Addressing consumer concerns/complaints.



BY BUILDING CUSTOMIZABLE, VISUALLY RICH, INTERACTIVE DASHBOARDS FOR EASY INSIGHT ACCESS, OR PLUG AND PLAY MODULES IN AN INTEGRATED BUSINESS DASHBOARD

AQ, with its state-of-the-art data visualization, can present analysis as a web-based dashboard or as a module belonging to a larger business performance study. It can, on the one hand provide users with a high-level summary of top insights and on the other, drill down to specific insights and data sets. Our dashboards help users consume data in an easy and intuitive manner and empowers them to derive insights quickly and meaningfully.

WITH OUR UNIQUE, HOLISTIC APPROACH TO TEXT ANALYTICS WE HAVE ENABLED SUCCESSFUL COURSE CORRECTION AND PROVEN POSITIVE BUSINESS IMPACT FOR CLIENTS ACROSS DIVERSE INDUSTRIES

WE HAVE HELPED



A popular US coffee and coffeemaker manufacturer



A leading cruise company



The largest automobile manufacturer in the world



A leading US pizza brand



A leading global chocolate manufacturer



+ Many More leading global companies

...AND WE ARE WAITING TO PARTNER WITH YOU



Neerav Naik is a founder member of Analytics Quotient (a Kantar Group Company, since March 2016) and leads the “Products” division in the organisation. His areas of expertise lies in the areas of Portfolio Management, Brand Strategy, Consumer & Growth modelling, Social Media Analytics and Data Visualization. He also excels at managing and growing large client relationships.

CONTACT AQ AT

✉ connect@aqinsights.com

👤 Amanpreet Singh Sooin | +14074540788 | amanpreet@aqinsights.com