

# The Truth about Text Analytics and Sentiment Analysis

**A CLARABRIDGE WHITEPAPER**



CLARABRIDGE

# The Truth about Text Analytics and Sentiment Analysis

A CLARABRIDGE WHITEPAPER



CLARABRIDGE

## Contents

<b>The need for analytics</b>	2
Manual analysis	2
<b>Outmoded techniques</b>	3
Outdated technology-based solutions	3
<b>A Better Way:</b>	
Clarabridge NLP	5
<b>Step one:</b> Normalization	5
<b>Step two:</b> Part of speech tagging	6
<b>Step three:</b> Named entity recognition	7
<b>Step four:</b> Semantic recognition	7
<b>Step five:</b> Advanced linguistics	8
<b>Step six:</b> Precise extraction	9
<b>Content detection</b>	10
Traditional approaches	10
Manual coding	10
“Bag of words”	10
Simplistic sentence-level scoring	10
<b>Sentiment analysis</b>	11
Clarabridge sentiment	11
Eleven-point scale	11
Negation and conditional sentiment	11
Amplification	12
Context-dependent sentiment	12
Exception rules	12
<b>Categorization</b>	14
Vertical and horizontal templates	14
Rules-based categorization	14
Auto classification	15
Machine learning	15
The absolute truth	16



## The need for analytics

Organizations have two kinds of customer feedback data that they measure, store, and analyze: structured and unstructured data.

**Structured data** is information that is clearly defined and easy to report on. It is the kind of data that is generally found in a survey and can be organized in a spreadsheet: name, location, age, and rating (3 out of 5 stars, for example, or a 10 for “most satisfied” versus a 1 for “least satisfied”).

**Unstructured data** as it exists today is, basically, text, although it can also include other media such as audio, photos, or videos. Unstructured data can be captured in an email, the “additional comments” section of a survey, voice recordings of customer interactions, a post on a customer review site, in social media, and dozens of other places.

Analyzing all this data correctly is critical, because it reveals everything from buying trends to product flaws and provides a significant business advantage. Organizations often struggle to do this analysis, however, because unstructured data is significantly harder to categorize and report on than structured data. It can be hard to parse due to grammatical errors or slang, it frequently contains multiple unrelated ideas, and it can represent various levels of sentiment related to each idea (for example, “I absolutely loved the food, but the waiter was rude and finding parking was impossible.”).

Unstructured data is harder to navigate

than structured data, but businesses must try—because **95% of customer feedback data is unstructured** and it has a wealth of information to empower your customer experience improvements. Overwhelming as it is, this kind of feedback is predicted to continue to grow, as millions of people increase their use of online resources. Unstructured content comes from emails, call center notes, insurance claims, contracts, press releases, research papers, forms, filings, medical records, survey responses, chat transcripts, voice transcripts, and from social and consumer review sites.

Organizations are drowning in this data, but still thirst for meaning; text can be a flood that threatens to swallow an organization that isn’t prepared to handle it. On the other hand, when it is understood and turned into actionable insights, it can be a treasure trove of customer, product, competitive, research, and marketing information. Organizations are able to better serve customers, control costs and risk, compete more effectively, and drive profitability. That’s why so many companies have explored so many different approaches to extracting these valuable insights.

Many approaches to text analytics have been tried and found lacking.

### MANUAL ANALYSIS

The most basic way to understand feedback text is simply to have someone read the text, note the contents, and categorize it. Market research-



## Outmoded techniques

ers, for example, often categorize, or “code,” the free-text responses in surveys. If an organization is only receiving a few dozen surveys a month, manual review and coding could suffice. However, this option doesn’t scale well for several reasons:

- **Cost:** Assuming minimum wage and that the average person can process 50 items of unstructured data an hour, it costs \$145,000 to have someone read through and categorize text for one million items.<sup>1</sup>
- **Time:** Using the same example listed above, one person can process 400 unstructured data elements in an eight-hour period. Going through one million comments would take 2,500 days to process, <sup>2</sup> or just under seven years. (To put this in context, consider that one Clarabridge customer processes over 700,000 post-call survey transcripts per month, which is just one of over 60 data sources they are collecting and analyzing. Yelp users post 37,987,200 reviews per day. TripAdvisor has 200 million posts and counting.)
- **Errors:** People make mistakes and are less accurate at coding than one may expect. Studies suggest that humans have an average accuracy rate of just 80%.
- **Inflexibility:** Manual processes can’t easily accommodate new categories or codes, and are even less likely to include “why” and “what if” analysis.

When these challenges are combined with the growing volume of unstructured data, as well as

multiple languages, it is easy to understand why automated processes have been explored for text analysis.

### OUTDATED TECHNOLOGY-BASED SOLUTIONS

While many of the earlier text analytics techniques are still in use, all have major flaws which ultimately have rendered them ineffective for true customer experience management.

**“Search engine” techniques** use keyword searches to find out which subjects are being discussed.

- **Overly simplistic interface:** Search tools were designed to be easy to use, which restricts their analysis capability to simple Boolean (not/and/or) expressions.
- **No insights:** Although a search tool may be great at rapidly returning documents, a user is still required to take the time to read through the returned documents to extract meaning from them. Because keywords can mean different things in different contexts, the tool may get many irrelevant responses while at the same time excluding relevant responses.

**Statistical techniques** search through feedback records and rank them in order of relevance based on the frequency of the query terms.

- **Retrieval inefficiency:** Lengthy documents are poorly represented because they have

<sup>1</sup> Assumes 1,000,000 pieces of unstructured data, processed at a rate of 400 a day. It would take 2,500 days for one person to process all of the unstructured data, working 8 hours a day at \$7.25/hour.



poor “similarity values” – in other words, they cover too many different topics in a single record that the models can’t account for.

- **Poor accuracy:** Documents about the same topic but which use different terminology are not captured, resulting in a “false negative match.”
- **Linguistic gaps:** The statistical approach lacks the structure to express important linguistic features such as phrases. To get a match, the queries have to contain many words to improve retrieval performance.

**Latent semantic indexing (LSI)** combines statistical and Artificial Intelligence techniques to find the latent structure in the pattern of word usage.

- **Inflexibility:** Although LSI techniques showed initial promise, they fail to factor in the ambiguities in human language.
- **Limited scope:** Long documents are still poorly represented, so relevant relationships may be hidden in masses of irrelevant findings or missed altogether.
- **Resource intensive:** LSI requires significant cost in terms of storage and computing time.
- **Incorrect findings:** LSI sometimes draws the wrong conclusions in terms of word relationships:

**Comment: The bed was extremely comfortable and the internet was slightly too slow.**

**LSI association results:**

BED -> SLOW

INTERNET -> UNCOMFORTABLE

UNCOMFORTABLE -> SLIGHTLY

## FIRST-GENERATION NATURAL LANGUAGE PROCESSING (NLP)

There are a variety of NLP techniques, leveraging different methods of parsing and interpreting the linguistic (words, grammar) and semantic (meaning) information.

**Shallow parsing** analyzes a sentence and identifies its constituents (noun groups, verbs, verb groups, etc.), but does not specify their internal structure nor their role in the main sentence.

- **Limited scope:** The simplified design of the shallow parsers reduces accuracy; for example, they cannot handle sentences with the main verb “be” or handle negative or passive sentences.

The **Triples** method looks for the appearance of three-word or three-phrase combinations in a sentence.

- **Excessive precision:** A Triples NLP engine only categorizes sentences that contain all three words in a pre-defined Triple. If only one or two of the words appear, the phrase is ignored.
- **Incomplete results:** Statistically, only about 50% of all comments will be recognized and categorized using the Triples method; manual or other processing techniques are required to account for the other half of the data.



## A Better Way: Clarabridge NLP

Recognizing the limitations of other methods, Clarabridge has developed a proprietary NLP engine using both linguistic and statistical algorithms, taking the best capabilities of each approach. The hybrid framework makes it easy to use without extensive understanding of statistics, comprehensive domain knowledge, or linguistic expertise.

The following sections give a step-by-step guide to how the Clarabridge NLP engine allows users to analyze a broad array of data sources with minimal fine-tuning while maintaining high standards of accuracy and coverage.

### STEP ONE: NORMALIZATION

Clarabridge NLP uses a workflow-like process to extract information from the text in a process known as normalization.

First, it separates the text into fragments for efficient information extraction. It identifies markers for the ends of sentences, paragraphs, and documents, and then breaks the stream of text into meaningful elements (words, phrases, and symbols).

Let's use a real-life example:

**My smart phone doesn't hold a chagre. Any suggestions**

Notice the typos and other grammatical issues—these are common in written language, so it

is important to have mechanisms in place to account for them.

Here's the first output the NLP agent produces:

My	smart	phone	doesn't
hold	a	chagre	.
any	suggestions		

Next, certain special characters and character sequences (like contractions and abbreviations) are removed and replaced with the appropriate words. In this case, the engine recognizes that “doesn't” equals “does not.” So the output is updated as:

My	smart	phone	<b>does</b>
<b>not</b>	hold	a	chagre
.	any	suggestions	

Sentence detection follows, where the input is clearly demarcated as two distinct sentences: “My smart phone does not hold a chagre,” and “Any suggestions.”

After sentence detection is **morphology**—which is the identification, analysis, and description of words, parts of speech, contexts, and other meaningful pieces of language (called “morphemes”).



The morphology module uses lexicons (pre-determined lists of words) and semantic dictionaries to identify the root form of each word (for example, the root form of “does” is “do”) and assign it a semantic category. Let’s go back to our example to illustrate what that means:

		Noun	Verb or auxiliary verb
My	smart	phone	do
not	hold	a	chagre
Negator	Verb, present tense		

Some entities may have more than one potential part of speech or other ambiguous value; these ambiguities are resolved in subsequent steps by the NLP engine.

The next step in the normalization process is undertaken by the Gazateer module. It takes the basic attributes assigned by the morphology module and adds additional information using standard lexicons, custom lexicons, and custom rules. This is the stage in which the tool understands that “smart” and “phone” become “smartphone.”

Next, the **unknown words module** uses a variety of methods including suffix patterns and common misspellings to try to identify unknown words. In this case, “chagre” becomes “charge.”

	“smart” and “phone” become “smartphone”		
My	smartphone		do
not	hold	a	charge
			“chagre” becomes “charge”

Finally, the **special entities dictionary** identifies specific types of entities such as addresses, dates, or financial information through a rules engine. It also includes a dictionary of idiomatic expressions that in English includes introductory phrases, complex prepositions, and complex adverbs. Clarabridge provides over 40 out-of-the-box dictionaries which contain domain-specific lexicons and thousands of rules that analyze the written data in ways that parallel the way humans process text. These dictionaries and rules are designed to be straightforward for a business user to customize; as a result, additional entities can be identified and new rules can be added.

## STEP TWO: PART OF SPEECH TAGGING

Next, the Clarabridge NLP engine detects parts of speech for every word. This detection uses a hybrid model combining statistical and rules-based approaches.

- Meaning is assigned to each entity in the



feedback record through the morphological module, as part of the normalization process described in Step One, above.

- **Ambiguity is cleared up.** As alluded to in the previous section, words can have many potential meanings; for example, “kind” can be either a noun (“kind of sweater”) or an adjective (“she was kind to the puppy”). To address these situations, machine learning is applied to select the most likely part of speech.

### STEP THREE: NAMED ENTITY RECOGNITION

After tagging the parts of speech, Clarabridge’s NLP engine performs named entity recognition. Named entity recognition looks for the probability that a sequence of words is an actual name. For example, “New York” would be recognized as a named entity (place) instead of being seen as two separate tokens – “New” and “York.” This ability to automatically identify and classify entities is important, since it enables users to quickly find the most critical pieces of information within massive volumes of text.

Named entity recognition is accomplished using four toolsets within Clarabridge:

- **Out-of-the-box lexicons:** Clarabridge provides built-in lexicons delivered with many types of named entities including Person, Organization, Location, Address, Product, etc.
- **Linguistic patterns:** Named entities are frequently written in specific patterns, such as when two concurrent words are capitalized (New York). Clarabridge identifies these linguistic patterns and determines the likelihood

of the text being a named entity or not.

- **Customizable rules engine:** Based on business requirements, it is often necessary to configure rules around recently created named entities; for example, a new product, brand, company, series, and so on.
- **Machine learning:** Similar to parts of speech tagging, machine learning enables the models to be trained to more accurately categorize within a specific dataset or to add additional entity types.

### STEP FOUR: SEMANTIC RECOGNITION

The next action taken by the Clarabridge NLP engine is semantic recognition. Semantic recognition detects the relationships between words, phrases, sentences, and larger units of text. Understanding how words relate to each other is essential to establishing and extracting meaning from text. Let’s use an example to illustrate the importance of semantic recognition throughout this section:

**My new MP3 player doesn’t work with my old laptop.**

Semantic recognition starts with a shallow parsing module that identifies the sequences of words that belong to a linguistic phrase. For example, this step enables Clarabridge to understand that “with my laptop” is a prepositional phrase, and how the words relate to each other within that prepositional phrase. Special entities, such as dates, currencies, numbers, etc., are also recognized.





Below is our example sentence after shallow parsing:

Noun Phrase	<b>My new MP3 player</b>
Verb Phrase	<b>doesn't work</b>
Prepositional Phrase	<b>with my old laptop.</b>

After shallow parsing, the NLP engine performs a breakdown of all relationships within a sentence. This is done through **full syntax parsing**, which takes the linguistic phrase structure from the shallow parse and builds a structured representation of that sentence as seen in the diagram above.

Using the full syntax parse as input, the **semantic network builder** then builds a network of semantic relations (“associated words”) between entities and events to determine the relationships underlying the structure. Using templates on top of the linguistic relationships, the **semantic network builder** also creates additional associated words that span a larger context based on the semantic network- see more about this in Step Six, “Precise extraction,” on page 9.

By extracting relationships between words (like an adjective and a noun, or “New” and “MP3” player in this example), the full semantic parse enables Clarabridge to understand how the adjective and noun relate to each other. This capability is critical to understanding what’s driving the feedback text in the first place.

By recognizing associated words, the Clarabridge NLP engine ensures that all data is understood and analyzed in all supported languages. Note that any languages that are not currently supported by Clarabridge NLP can still be processed using Clarabridge using a tokenization method that provides the individual words needed for categorization rules.

## STEP FIVE: ADVANCED LINGUISTICS

The full semantic parse completed in step four enables advanced linguistics such as clause detection, negation support, and anaphora resolution.

**Clause Detection:** The clause detection feature does just that—detects unique clauses within the text. A clause is the smallest grammatical unit that can express a complete proposition (essentially an idea). Generally, clauses are subordinate portions of a sentence that consist of an independent thought within the sentence. Clarabridge’s NLP engine identifies these independent clauses so that their semantics and sentiment don’t interfere with one another. Let’s remember our previous example: “I liked the food, but the waiter was rude.” This sentence is comprised of two independent thoughts: one related to the food and the other related to the service. Clarabridge detects these as distinct clauses.

**Negation support:** Sentiment is often negated by words like “no,” “not,” and “never,” and sometimes the distance between the negator and the sentiment word can be quite large (5+ words away). To ensure that sentiment is accurate, the Clarabridge NLP engine provides negation support, which flips the sentiment on any positive or



negative word to which it is applied. For example, in the sentence, “I can’t give your store a good review,” “not” (recall that “can” and “not” are separated out in an earlier step) is applied to “good” which then becomes negative (correct) instead of positive (incorrect).

**Anaphora resolution:** Anaphora resolution enables the system to automatically resolve the pronoun “it” back to its proper noun. Also, any associated words that occurred with the pronoun will be updated to reflect the resolved noun. For example, Clarabridge will correctly determine that a car was gliding through the air in the sentence, “The car was really great; it [car] cruised along like it was gliding through the air.”

## STEP SIX: PRECISE EXTRACTION

The semantic network builder enables precise extraction, by overlaying a set of templates onto the semantic network. This allows the system to target entities that span more than a single relation so that a new associated word pair can be created. For example, in the sentence “Julie is tall,” the words “Julie” and “tall” do not have a direct relationship that can be analyzed. Instead, “Julie” has a relationship with the verb “is” (or “be” after the Normalization step), and so does “tall.” Precise extraction enables the system to recognize that the meaningful relationship is between “Julie” and “tall.”



## Content detection

Once text has gone through Clarabridge NLP, it is ready for the next stage of analysis: content detection. This means determining if a feedback record is really customer feedback and if it expresses sentiment regarding the brand. If it meets those criteria, it is said to be “contentful.” This is in contrast to “non-contentful” data, which does not contain the kinds of information that can reveal customer insights. Non-contentful data includes spam, advertising, and other text that may include your key terms but is not interesting for analysis.

The separation of contentful and non-contentful data is done using sophisticated statistical models that recognize the specific linguistic qualities that characterize both types of data. Non-contentful data is distinguished by traits including certain word combinations, patterns of punctuation, and the way capitalization is used.

It is critical to remove non-contentful data from your analysis prior to reporting because it can skew results.

For example, see the following two tweets:

**I'm SOLD! Loving that the @MakeupBrandX Nailpolish requires no base coat and looks so pretty in natural light!**

**Compared to:**

**@MakeupBrandX Nailpolish:  
Buy 1 Get 1 Free  
@CosmeticsStore**

A simple statistical analysis would show that both of these feedback records include references to Makeup Brand X—however, only the first one has customer insights available. Beyond the out-of-the-box content detection functionality, the Clarabridge content detection mechanism can also be retrained using customer data for very specific results.



## Sentiment analysis

After content detection, the next step in text analysis is to determine sentiment—the positivity or negativity expressed through the text.

As with language processing, there are multiple ways to handle sentiment analysis. These vary with respect to both efficiency and accuracy.

### TRADITIONAL APPROACHES

**Manual coding** is the process by which a person simply reads a document and assigns it a sentiment value—usually “positive” or “negative.” As with the manual approach to analysis discussed above, this method is time-consuming, imprecise, subject to human error, and does not scale to large amounts of feedback.

**“Bag of words”** is a method that looks at the feedback document and assigns positive, negative, or neutral values to each word based on a pre-defined, built-in dictionary. Drawbacks include:

- **Inaccuracy:** Just looking at a group of words is very different from looking at how those words relate to each other.
- **Ambiguity:** Simply detecting “emotion” words may indicate whether a sentence is positive or negative; however, it does not explain the topic to which that sentiment is directed.

**Simplistic sentence-level scoring** is the way that most solutions categorize and provide sentiment analysis, by not digging deeply enough into the entire context. The problem with this is Imprecision. For example, in the sample sen-

tence, “I liked the food, but the waiter was rude,” sentence-level scoring would be likely assign this sentence a positive sentiment value based on the first clause. It misses out on the complete picture by being unable to separate and then separately analyze every unique idea, thought, and phrase.

### CLARABRIDGE SENTIMENT

Clarabridge overcomes the limitations of other sentiment analysis techniques by combining lexical and grammatical approaches. The system understands how sentiment is altered or intensified by taking into account grammatical constructs and function words. This gives Clarabridge the unique ability to accurately handle negation, conditional sentiment, domain-specific sentiment, and exception rules.

### ELEVEN-POINT SCALE

Clarabridge indexes the sentiment score on a normalized minus five (-5) to plus five (+5) scale. This provides significantly greater accuracy over the “positive,” “negative,” or “neutral” rating—since we all know that there is a big difference between a good meal (+1) and the best food you’ve ever eaten (+5). For further analysis, users can perform sentiment filtering to isolate the degree of positivity/negativity, and root cause analysis to determine what is driving the sentiment.

### NEGATION AND CONDITIONAL SENTIMENT

**Negation** and **conditional** sentiment are two ways in which the relationships among the words



in a sentence can impact the overall sentiment.

### Negation:

**I am not happy about the repair experience.**

The word “happy” would typically carry positive sentiment. However, because it is paired with “not” in this sentence, the positive sentiment is negated, which Clarabridge recognizes.

### Conditional sentiment:

Similar to negation, conditional sentiment refers to statements that may seem positive but are neutralized by phrases such as “kind of” and “sort of.” This is calculated using positional rules, which identify when these qualifying phrases are used together. Other examples of conditional sentiment include:

#### **COULD + HAVE + BEEN + [positive] = negative**

Example: It could have been a good experience

#### **COULD + HAVE + BEEN + [BAD/PROBLEM] = neutral**

Example: It could have been a problem, but...

#### **USED + TO + BE + [positive] = negative**

Example: The restaurant used to be cleaner.

### AMPLIFICATION

A positive (or negative) phrase can be made more positive (or negative) by modifiers. Clarabridge recognizes the difference these modifiers make. For example, “The food was great,” (+2 positive sentiment) jumps to a +3 positive when written as, “The food was so amazingly great.”

### CONTEXT-DEPENDENT SENTIMENT

Context-dependent sentiment is the ability to recognize that in some contexts a phrase may be positive, whereas in other contexts it is negative. One example of this is the word “thin.” In technology, “thin” things are generally considered good – think of a thin laptop. However, in the hospitality industry, “thin sheets” or “thin walls” are a negative. Clarabridge uses pre-defined rules and a taxonomy to get this fine-tuning underway quickly. Examples of industry templates include retail, automotive, high tech, financial services, and restaurants; horizontal templates include customer service, online experience, cries for help, and social media.

For individual situations, the Clarabridge platform also enables word-level tuning, the ability for the pre-defined sentiment attached to a specific word to be increased or decreased based on the particular need. This process is done easily, without the need for technical support or linguistic expertise. Sentiment tuning can be applied universally or only for a specific project or sub-set of data.

### EXCEPTION RULES

Exception rules are pre-defined rules to account for different linguistic constructs that consistently change the sentiment of a word or phrase.



Consider the following sentence:

**The shirts were  
too orange.**

The color “orange” would normally have a neutral sentiment value (a “0” on the 11-point scale). However, adding the word “too” changes the sentiment to a negative. This holds true often enough that it is valuable to have a rule within the tool stating that  $TOO + [NEUTRAL] = NEGATIVE$ . Each sentiment exception rule can also be turned on and off as needed (customizable).

The Clarabridge platform includes over 500 built-in exception rules for English processing, and also enables users to easily add their own rules to meet their specific needs.



## Categorization

After the text has been processed for underlying structure and sentiment, Clarabridge uses both linguistic and statistical techniques to group the data into related buckets. The goal of the process is to categorize as much of the data as possible (known as “high recall”) and as accurately as possible (known as “precision”). This enables easy reporting and promotes discovery of insights in the data. Clarabridge conducts categorization with unique features including templates, rules-based categorization, auto classification, and machine learning.

### VERTICAL AND HORIZONTAL TEMPLATES

Templates, discussed in the sentiment section above, provide additional benefits beyond sharpening sentiment analysis; they also help to categorize textual data more quickly. One, none, or multiple templates can be applied to the same data set – allowing deep-dive analysis in specific areas, such as the online shopping experience for cell phones. Further, these templates can be extended and customized based on business needs; users can build their own models as well.

In the example to the right using the Lodging Template, all items related to lodging are automatically classified into groups: bell staff, bathroom, check-in/out, concierge, etc. Within those sub-groups, the content is

further refined to additional categories. All of this organization is done without manual intervention.

### RULES-BASED CATEGORIZATION

In addition to the out-of-the-box templates, Clarabridge users can also add specific categorization rules. All extracted entities and relationships are displayed, so that business users can quickly refine the categorization model. Words can be dragged and then dropped into the definition of the category, bringing along any other words based on the word stem. For example, if a category was defined to incorporate the notion of “clean,” the solution also identifies instances of “cleaning,” “cleaned,” and “cleaner” within the data set, as well as misspelled instances such as “claen.”



## AUTO CLASSIFICATION

Even after the built-in templates and rules-based engines categorize the data, there will still be some text left which doesn't naturally fit into the buckets created. Auto classification, also known as the "other" bucket, is designed to handle this left-over text. Leveraging advanced statistical algorithms, auto classification automatically builds category trees for that "other" data. As a result, users can load millions of different pieces of text and Clarabridge will quickly sort through it – basically building an outline without prompting, so you can understand what's in the "other" bucket and put it in the appropriate category going forward.

## MACHINE LEARNING

The Clarabridge system is able to refine its categories over time based on rules created and data already categorized. Just like the machine learning that takes place during parts of speech classification, the solution understands what categories "look like" based on past experience. For example, the sentence below:

**I couldn't stand the clanking of the A/C in my hotel room.**

The example clearly deals with unacceptable noise levels caused by the air conditioning in a hotel room, even though there is no mention of "noise" in the textual comment. Based on previous experience (in this case, that "clanking" and "noise" are commonly connected linguistically

in other examples of the text), machine learning would understand that this comment is actually about noise and categorize it appropriately. As a result, the need to manually code every document is eliminated, since the system is able to expand coverage of categories over time without any intervention.





## The Absolute Truth

The Clarabridge approach to processing text is different from other methods. As outlined above, it includes a variety of components supported by very sophisticated technology. Why does this matter? **Because better information leads to better decisions.**

Clarabridge technology is designed to maximize the accuracy of the NLP, categorization, and sentiment analysis. As a result of the detailed linguistic and technological methodology detailed here, Clarabridge users are able to avoid false negatives (or false positives) when drawing actionable insights from their unstructured data. When you can rely on your insights, you can act confidently to improve the customer experience.

The accuracy and intricacy of Clarabridge text analytics provides a framework for the entire Clarabridge suite. With advanced visualizations, solutions for sharing and collaboration, and more, Clarabridge technology empowers organizations to process the growing expanse of customer feedback accurately, efficiently, and with a purpose – extracting actionable insights to inspire action.





CLARABRIDGE

# About Clarabridge

## OUR MISSION

Our mission is to help businesses win the hearts of their customers through emotional intelligence. An emotionally intelligent organization will develop lasting and positive relationships with its customers that transcend momentary challenges and threats.

## FAST FACTS

- SaaS provider of customer intelligence and analytics
- 450+ global brands served
- Founded in 2006
- Headquartered in the Washington DC metro area with offices in San Francisco, London, Barcelona, and Singapore
- Led by CEO and founder Sid Banerjee, named as *Washington Business Journal's* most admired CEO in 2014
- 250+ employees worldwide
- Offers the world's most comprehensive customer intelligence platform, powering customer experience management programs
- Served customer experience professionals, marketers, customer care leaders, and operations managers
- Key clients include ADP, Cisco, Dell, G.E. Healthcare, Orbitz, PetSmart, Red Roof Inn, Unilever, and Verizon

## OUR SOLUTION

Clarabridge helps you find, understand and use customer feedback that is hidden in silos across your organization and in other public forums. Putting customer intelligence to work empowers your business to make big and small decisions and drive your customer experience management programs.

Clarabridge offers the world's most sophisticated customer intelligence platform. This platform connects to all sources of customer feedback and analyzes it to detect emotion, context, and root causes, as well as predict future customer behaviors. The analysis is presented in dashboards and reports and can be used for driving action with customer engagement, case management, and alerts.

## WHY CLARABRIDGE?

### **Connect to every source of customer feedback.**

Clarabridge is the only technology platform that can analyze feedback data from all sources, all in one centralized hub. Clarabridge analyzes data from social media, online chat, call center recordings, agent notes, surveys, online review sites, and many other sources.

### **Smash silos, and empower your entire organization.**

See customer feedback come to life in user-friendly dashboards, reports, and alerts. Share this information through your entire business, and create a customer-centric culture.

### **Tune the solution to meet your specific business needs.**

Even though we offer out-of-the box industry templates to get you started, you can tune the system to your exact business needs to ensure that the information is relevant and actionable.