

Integrating customer experience management with marketing research

Published: [January 2014](#)

Author: [Rick Kieser](#)

Article Abstract

From natural language processing to machine learning, Rick Kieser gives an overview of various approaches to processing the many forms of customer feedback now available to companies and makes a case for a multi-modal approach.

Editor's note: Rick Kieser CEO of Ascribe, a Cincinnati software firm.

Thanks to ever-expanding options for digital communication, customers now have countless ways to share their feedback with companies – and each other – with both positive and negative messages. Many customer-oriented businesses now have customer experience management or voice-of-the-customer platforms in place to manage this heavy inbound flow of communication.

In parallel, companies continue to solicit insights through traditional surveys, though some are now questioning this approach. The same advances in digital communication have spawned alternative knowledge providers who claim to be able to provide better insight from this raw material than old-style market research.

It is rare for these disciplines to be integrated or for market researchers to use the feedback obtained from these operational platforms as a primary source for insight generation. In this article I will argue why an integrated approach is a necessity. Without it, market research as we know it will be marginalized. I will also attempt to show how recent advancements in processing unstructured text can put this strategy well within the reach of market researchers.

Virtually no difference

Customer experience is increasingly the only meaningful differentiator for the consumer. Rigorous attention to product design and quality assurance today means there is virtually no difference in the quality or effectiveness between the brands found in most product sectors. The speed and transparency of the online marketplace means consumer decisions today – even about offline purchases – are all about price, past experience and reputation, judged from others' experiences.

Can we put a value on the importance of reputation? According to one survey, more than \$83 billion in business is lost each year due to poor customer experiences¹. In this dynamic and rapidly changing environment, marketers must constantly reexamine how to understand and engage customers, re-prioritize channels and maintain a consistent customer experience across all channels, while optimizing their marketing spend.

Each initiative can only be achieved with reliable data used to generate high-quality insight with which to drive decision-making. In the past, market research was often the only viable means to gather and process this information. Today, there is usually an abundance of data on every different aspect of the customer experience and it can often be integrated easily. Now it is the market research data that can appear out of touch and disconnected with business.

The customer experience management and voice-of-the-customer (VOC)² industries were born approximately 20 years ago and although originally distinct, today the terms seem largely synonymous.

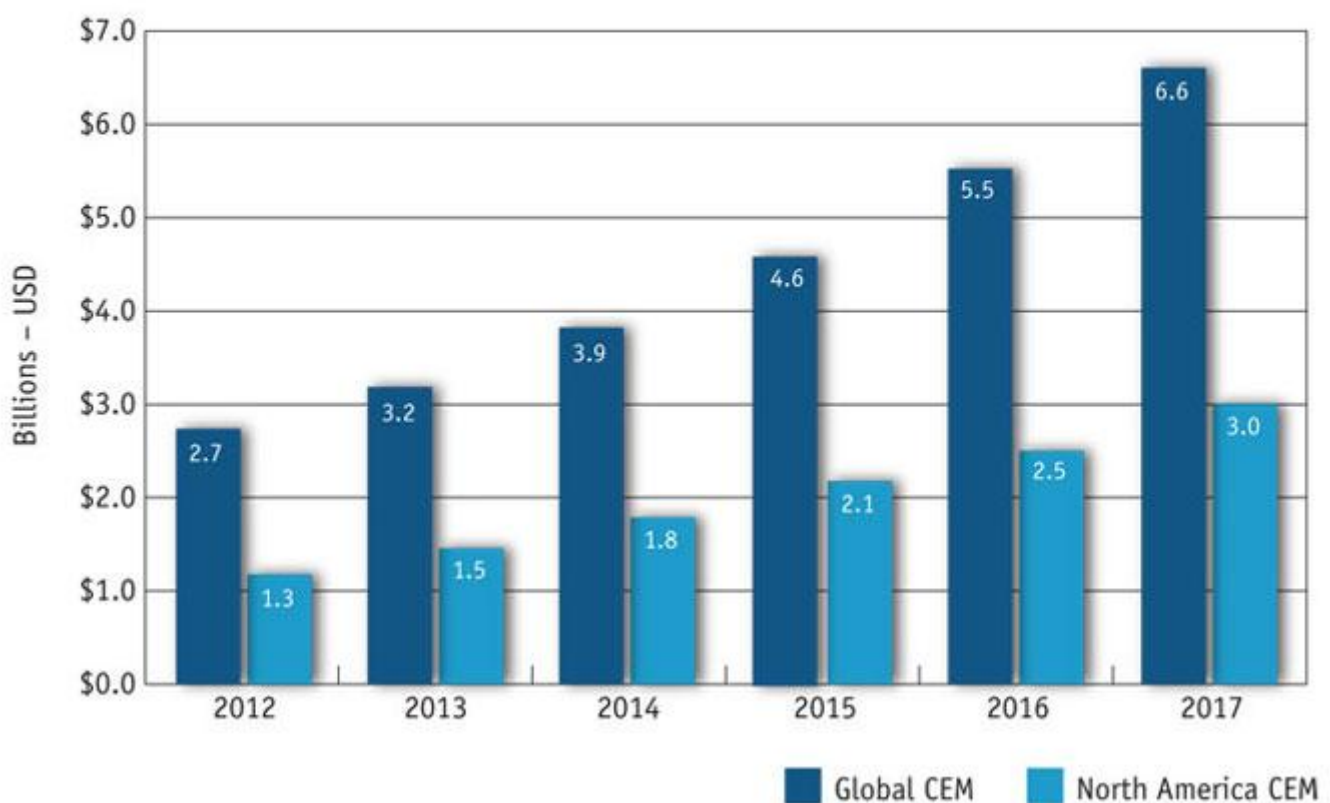
Both focus on using automated processes to track, monitor and organize every interaction between a customer and the organization throughout the customer life cycle.

In parallel, enterprise feedback management (EFM) is a term coined by research technology providers to describe survey-based feedback solicitation. An EFM interaction is usually triggered by a customer touchpoint, which makes it a convenient data source for many VOC initiatives. But some VOC initiatives do not depend on surveys and capture their data elsewhere.

The CEM industry has been growing nearly 20 percent annually and the market is projected to nearly triple between 2012 and 2017 (Figure 1) to more than \$6.8 billion³. This is in contrast to market research, which stands at \$39 billion of revenue worldwide, and inflation-adjusted growth of 0.7 percent achieved in 2012⁴.

Figure 1

Customer Experience Management (CEM) Market



I see three principal driving forces behind the growth of the CEM industry:

1. Enterprises have become more customer-centric, realizing the closer they are to meeting their customers' expectations, the better chance they have of succeeding.
2. Companies tend to understand that acquiring a new customer costs more than keeping an existing one (one source states up to five times as much⁵).
3. Technology now makes it easy to track customer interactions across a wide range of customer touchpoints and generate insight from this.

This growth has resulted in a tsunami of customer feedback, which must be transformed into insight if companies intend to use it. It is the new raw commodity in every business: Processed and applied in the right way, data holds tremendous value. But with no effort made to generate insight from it, this data simply accumulates without contributing to the company's progress.

Do not control

The old paradigm for customer feedback in both market research and EFM relied on survey-based interventions, directed and regulated by the company. Initially, the rise of digital communications simply added more channels to the solicited model, with the advent of online surveys and communities (MROCs). But now, with rise of social media, consumers have become voluble through channels that marketers do not control. For example, customers share their experiences with the world through the likes of Facebook, Twitter, review sites like TripAdvisor and personal blogs, without even being asked.

This abundant unsolicited feedback challenges traditional market research approaches, with their controlled samples and carefully constructed questions. This creates a new competitive environment based on advanced analytic technologies to analyze these vast amounts of unstructured numeric and textual data. It has encouraged new entrants, in the guise of data providers and knowledge-management companies, who are very different in their approach from research companies.

It is not only this reversal of flow in insight gathering that is difficult; coping with the volume of data and its highly unstructured nature is one of the greatest challenges of big data. Only technology can manage this massive inflow of data. And because so much of the content is raw text sitting alongside numeric data, we must use tools that can interpret both subjective and objective variables – both text and “hard” data.

A taxonomy of text insight technologies

Just as CEM software has reached a level of maturity, so too has software for text analytics, mostly based on an underlying set of techniques termed natural language processing (NLP). The predominance of this method in commercial software risks overshadowing two complementary text processing methods, which are equally relevant and in some instances are more appropriate in handling very large volumes of feedback data.

Natural language processing. The technology behind NLP or text analytics uses lexicons or dictionaries alongside a series of deterministic rules to bring together responses that appear to share similar content, to identify topics or to identify sentiment such as positivity or negativity. NLP is particularly well-suited to discovery of meaning or sentiment in large data sets, when used as a query or interrogation tool, as well as in developing analytical frameworks – an approach some describe as text mining.

To achieve acceptable levels of accuracy, highly-skilled human input is required to interpret what is discovered and to refine the process of topic and sentiment extraction by writing additional rules. This tuning process is time-consuming and expensive and is often overlooked when selecting this method.

Machine learning. A viable alternative to NLP is machine learning, an artificial-intelligence approach that learns how to categorize and interpret text automatically from a sample of training examples that have been coded manually. As more examples are subsequently provided, including any that arise from quality-control checks, overall accuracy improves.

Machine learning is especially well suited to large-scale repetitive tasks and can run as an automated process, once trained, with minimal intervention. Because a training set is required (typically 20 examples for each topic or category), the initial investment of effort lends itself to larger-scale or continuous projects; however, once operational, unlike NLP, it does not demand expensive interpretive or technical input on an ongoing basis, often making it a lower-cost automated solution when compared to NLP.

Semi-automated or computer-assisted coding. Auto-assisted methods organize the work intelligently and optimize human decision-making in classifying customer comments by using powerful searches, fuzzy

matching within an overall organizing structure. Arguably, this approach can yield the most accurate results but it does not scale well. The effort increases in line with the volume of work and the management burden increases exponentially. Beyond low volumes, this approach carries a high cost premium compared to other methods.

If a company is analyzing the same type of customer feedback from a single source, one technology may be all that is needed. However, in today's diverse and data-rich environment, that is rarely the case. Much of the challenge in handling multiple channels of feedback, both solicited and unsolicited, is in the diversity of their characteristics. An appropriate analogy is the bicycle. If the road to be traveled is flat and straight, a single-speed bike will be fine. However, in more varied terrain, a multi-speed bike will make the trip easier, faster and more efficient.

For example, a transcribed phone call will exhibit very different use of language compared to a Facebook update, as will comments in a customer survey differ from comments made on a hotel review site. Comments on Twitter are entirely different again, with their cryptic content and embedded cross-references. Each type of language demands a different kind of solution. In my experience, even large enterprises can struggle to find an effective solution, because there is a common misconception that a single technology (often assumed to be NLP) can apply to all situations.

Five considerations

To optimize customer insight there are five considerations that need to be balanced in any effective VOC or CEM program, namely: the program objectives; feedback channel types; insight technologies deployed; cost constraints; and volume of responses. Since all of these go beyond the scope of this article, I will focus on the main consideration. A combination of technologies is needed to facilitate effective feedback management across all channels. To understand which, you must know your channels as well as your text-insight technologies and then match those channels to the best combination of technologies.

Any enterprise is likely to have a wide range of feedback channels. Each, as described here, has very different characteristics and presents its own challenges from an analysis perspective.

Social media comments can exist in very high volumes but can be difficult to interpret, as they are often extremely context-specific, cryptic, jargon-riddled and even ironic. Nevertheless, social media is a source of insight that can help companies react quickly to events or new situations and reveal weak or emerging trends.

Survey data, on the other hand, is much easier to interpret, since questions are predetermined and answers are more focused. The high cost in obtaining survey data usually means the volume of data is low, so weak or emergent trends may be entirely absent from the data collected.

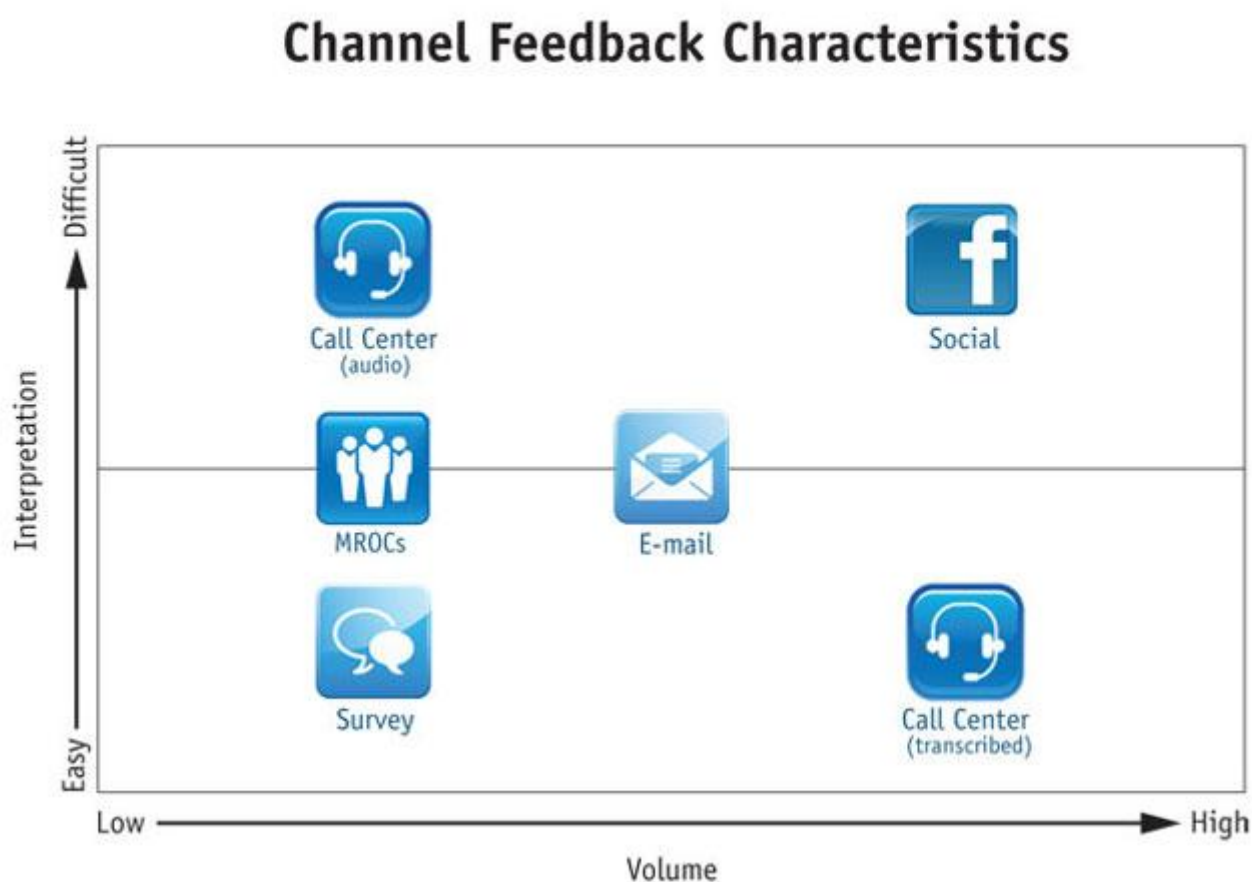
Call-center transcripts can offer a high volume of responses and are relatively easy to interpret, since they are intermediated, and may have been summarized by the call-center operator. Call-center audio, in comparison, can be harder to interpret and usually requires transcription first, making it a lot less accessible.

Market research online communities data is often largely text-based but relatively low in volume. Although it can superficially appear social media-like in its format, it is much easier to interpret because contributions are carefully considered by the contributor and tend to be focused and context-specific.

Inbound e-mail, in many ways, sits at the center of these channels in terms of its form and structure. Customer service e-mails can be highly structured, with a series of questions and answers not dissimilar to a survey, executed as an extended e-mail conversation.

In searching for the right tools and methods to apply to each channel, it is essential to understand the characteristics of the available channels within any specific enterprise. The two most relevant variables are the volume of data and the difficulty of interpreting the channel. Analyzing these will reveal important differences in the scale of the task to be undertaken. A generalized visual representation of this appears in Figure 2.

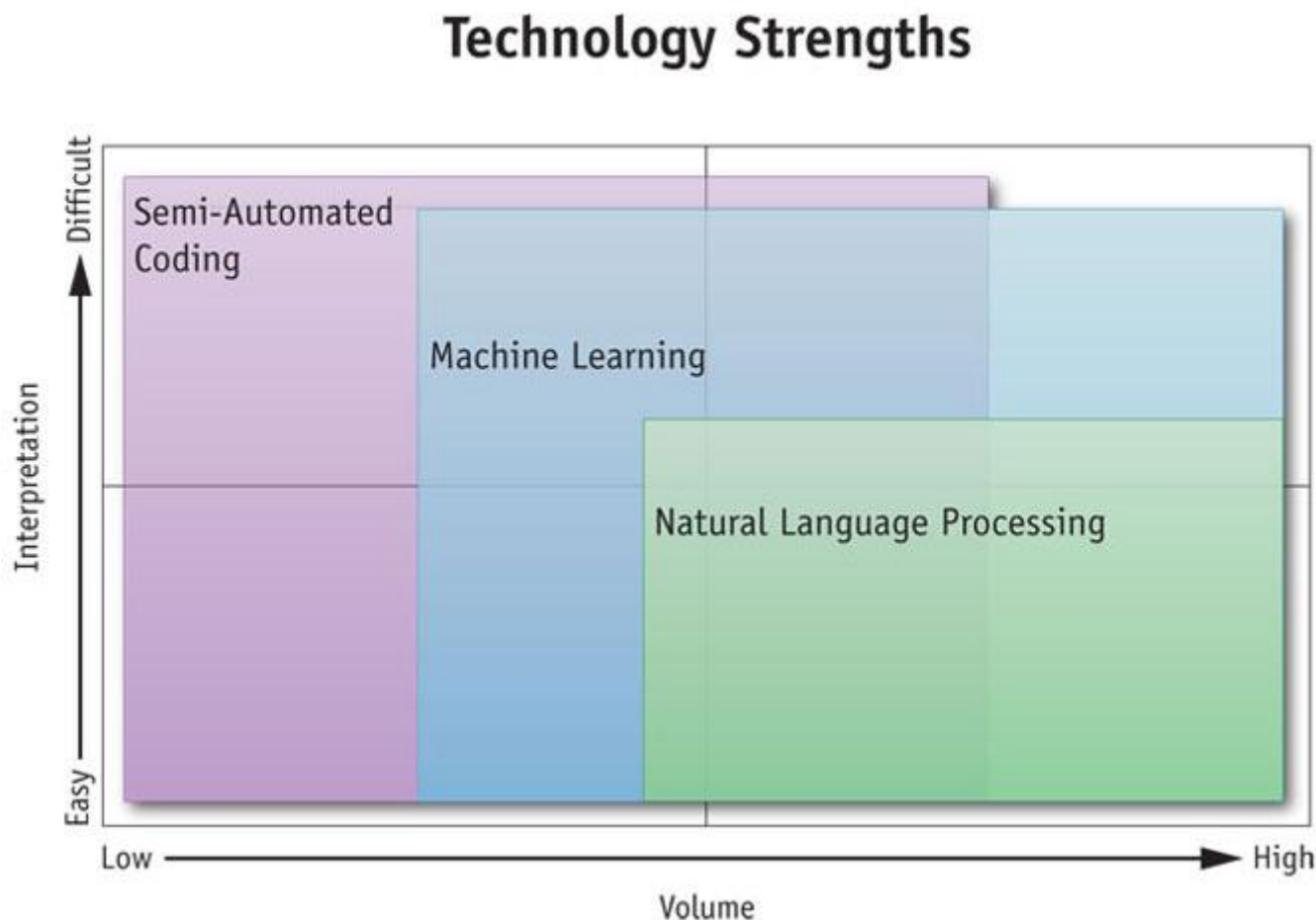
Figure 2



Match the technology to the feedback type

The final step is to match the best technology to the feedback type. Figure 3 shows the optimum reach, in terms of cost-effectiveness, for each of the three different technological approaches with regard to the same two variables: volume and difficulty of interpretation.

Figure 3



NLP can handle high volumes but because it is a mining tool more than an interpretive tool, it is better at providing fast directional feedback. It is less amenable to tuning to particular contexts and so it is less effective where there is a mix of channels that are more difficult to interpret.

Machine learning works best when handling a high volume of text, as the marginal cost of adding more work is negligible. Because it is trained using context-specific examples, it is capable of interpreting difficult-to-recognize concepts or emotions once it has an effective set of training examples.

Both automated methods require effort in applying them to a new channel, either in adjusting the rules or dictionaries for NLP or in training and validating for machine learning.

Semi-automated methods are well suited to channels where the content is difficult to interpret because humans are involved in determining meaning. Further, they do not have the same start-up threshold to cross as automated methods, making them cost-effective for low volumes. But, because little economy of scale is possible for high volumes, semi-automated methods tend to price themselves out of the equation for high-volume channels.

It may be possible to find an optimum solution if you are working with only a single channel. However, with three or four different channels, you certainly need more than a one-speed bike to cover the terrain (to return to my earlier metaphor). This is not an inherent weakness of automated text processing; rather, judiciously combining methods actually builds a more robust and more cost-effective solution that plays to the sweet-spots of each technology involved.

In practice, this means you can optimize not only time and quality but also reduce cost by concentrating human intervention within a few small areas where its value can be amplified – such as highly targeted quality control or having analysts interpret automated text-mining reports.

Do not settle

Set against the backdrop of different – and often disjointed – customer insight initiatives and customer feedback channels, companies can now start to build highly effective technological solutions to integrate feedback, provided they fully evaluate their internal needs and do not settle for suboptimal technological solutions. The technology now exists to facilitate transforming disparate feedback channels – both solicited and unsolicited – into timely, actionable insights. While some investment is inevitably required, the results can be extraordinarily valuable and a full payback should be expected within a single fiscal cycle.

According to the Harvard Business School, increasing customer retention rates by as little as five percentage points has been shown to increase profits by between 25 and 95 percent⁶. The investment in the technologies I have discussed here are likely to be a small fraction of one percent of revenues.

The good news is that the best is yet to come. The technology continues to improve. Companies now have a choice of effective tools they can use in combination to manage their incoming flow of customer feedback. Market researchers have an even richer source of data available to them to complement survey data and more conventional hard number big data feeds. In a few years, I am sure we will look back on today and be surprised that we only managed to unlock a fraction of the potential insights waiting to be discovered.

References

¹ The State of Marketing 2013, IBM's Global Survey of Marketers (2010-11 Research Study by Datamonitor/Ovum).

² "VOC" term was first used in 1993 within a paper written for *Marketing Science*, authored by Abbie Griffin and John Hauser. (University of Chicago and Massachusetts Institute of Technology).

³ Customer Experience Management & Voice of the Customer Analytics Market Size and Forecasts (2012-2017), by MarketandMarkets, 12/2012.

⁴ ESOMAR Global Market Research Report 2013.

⁵ Confrimint Community Conference, keynote address by Henning Hansen, CEO, Las Vegas, 2013.

⁶ "The economics of e-loyalty," Frederick F. Reichheld and Phil Schefter, *The Harvard Business Review*