



Text Analytics:

Practical examples for
Using Voice of the Customer to
improve your operations

FOREWORD

Big Data, predictive analytics, and advanced analytics are all very hot topics right now for every company. A growing percentage of that big data is in the form of text – and companies are starting to evaluate how to incorporate the analysis of the text data into their full-scale analytics initiatives.

We [published an e-book](#) where 28 industry experts provided advice on how to get business value from text analytics. We found that a few key themes emerged from the essays in the book:

- 1) There are a multitude of areas where text analytics can be useful to an organization. When selecting how and where to use text analysis, it's important to think outside the box to all areas of your business – beyond just “how do my customers feel about my brand”.
- 2) When starting to incorporate text analytics - pick one or two problems or areas to begin with, rather than trying to start out with a full-blown organizational initiative. It would most likely fail. Do a pilot - get an executive sponsor and a stakeholder group, work through it, and document lessons learned and results. Then continue to move to the next problem.
- 3) Regardless of the technology available for analyzing texts, most of our contributors will agree you cannot get rid of the human element in analyzing text data. So don't try to!

We've selected a sub-set of these essays to demonstrate how specifically, the voice of the customer via text mining can be valuable across many operational areas. We hope these essays get your wheels turning about how text analysis can benefit your bottom line.



Vicki Rollins
VP Marketing
Angoss Software Corporation

About Angoss Software

Angoss is a global leader in delivering predictive analytics to businesses looking to improve performance across risk, marketing and sales. With a suite of big data analytics software solutions and consulting services, Angoss delivers powerful approaches that provide you with a competitive advantage by turning your information into actionable business decisions.

Many of the world's leading organizations in financial services, insurance, retail and high tech rely on Angoss to grow revenue, increase sales productivity and improve marketing effectiveness while reducing risk and cost. Angoss serves customers in over 30 countries worldwide.

INTRODUCTION

Two great forces are converging on businesses all around the world. One is a tidal wave of unstructured data in the form of text, audio, images, and sensor inputs. The other is a whole new generation of data processing technology, including low-cost, scalable cloud storage of almost unlimited size, and new techniques for quickly analyzing unstructured data. The result is an explosive growth in knowledge and insight.

Although analyzing text for insight is not new, what has changed in recent years is the ability to mine vast quantities of text—such as all the content on the Internet—and to do it quickly. This capability is profoundly changing how businesses use information to learn about markets, trade on knowledge, and refine their operations. Yet text analytics methods and techniques are rapidly changing. So, what are the best ways to extract value from text? With the generous support of Angoss, we posed the following question to 28 text analytics experts:

What advice would you give someone in your industry to get business value from text analytics?

The responses we received reflect the vibrant and evolving state of this emerging technology. One startling revelation that jumped out at me as I read these articles is that just as machine learning speeds the breadth and depth of analytical insight, machine-driven text analytics is having an extraordinary impact on the speed of human learning.

Even if you are not currently involved in text analysis, you cannot help but feel captivated by the insights this e-book contains.



All the best,
David Rogelberg
Publisher



Mighty Guides make you stronger.

These authoritative and diverse guides provide a full view of a topic. They help you explore, compare, and contrast a variety of viewpoints so that you can determine what will work best for you. Reading a Mighty Guide is kind of like having your own team of experts. Each heartfelt and sincere piece of advice in this guide sits right next to the contributor's name, biography, and links so that you can learn more about their work. This background information gives you the proper context for each expert's independent perspective.

Credible advice from top experts helps you make strong decisions. Strong decisions make you mighty.

Practical examples for Using Voice of the Customer to improve your operations



Roman Kubiak
Senior Consultant.....5



Thomas Kunjappu
Twitter.....7



Meta Brown
A4A Brown, Inc.....9



Bill Sheldon
Angoss.....11



Kaenan Hertz
Ernst & Young.....13



Shree Dandekar
Dell Inc.....15



Ramkumar Ravichandran
Visa Inc.....17



Seth Earley
Earley Information Science.....19



Sudip Chakraborty
Xypress LLC.....21



Daniel Angus
The University of Queensland....23



Helen Clegg
A.T. Kearney.....25



Jeff Catlin
Lexalytics, Inc.....27

COMBINE STRUCTURED AND UNSTRUCTURED DATA TO IMPROVE CUSTOMER RELATIONSHIPS



**ROMAN
KUBIAK**

Senior Consultant

Roman Kubiak holds an M.S. degree in economics from Gdansk University and an MBA from Baruch College. He has worked on analytic teams for such financial services firms as American Express and CityCards as well as for e-commerce companies 1800Flowers.com and PetCareRx.com. He moved to a telecommunications company to pursue his interests in natural language processing and text analytics. His other interests include online marketing attribution and net lift purchase propensity modeling.



Twitter | Website



Download the full e-book: [Text Analytics](#)

For Roman Kubiak, senior consultant to the advanced analytics team at a telecommunications company, text analytics helps businesses learn more about their customers. We derive the most value, he says, by mining both structured and unstructured data. “Today, many of the interactions are long distance. People are either interacting with our website or exposed to our ads; they see something on the Web. Other than in retail stores, we don’t have much of an opportunity for face-to-face interaction,” he says. “Instead, we have the customer service phone transcripts, the customer support emails, and what people are saying on social media, and we use that information to learn what customers’ impressions of our business are.”

It’s not a simple process. Kubiak says that he uses tools like topic modeling, text classification, and text tagging to merge the output from unstructured data with the output they receive from numeric databases and customer records. “In my experience, the biggest challenge is capturing and storing the text data generated by customers,” he says. He recommends doing so as a good starting place. “You need to be able to keep records for extended periods to be able to analyze changes over time.”

“ We have the customer service phone transcripts, the customer support emails . . . we use that information to learn what customers’ impressions of our business are. ” 

KEY LESSONS

- 1 The ability to capture customer data over an extended period of time is essential to developing a successful text analytics strategy.
- 2 Preprocessing data—cleansing, tagging, classifying them—is a time-consuming element of text analytics that cannot be overlooked.

COMBINE STRUCTURED AND UNSTRUCTURED DATA TO IMPROVE CUSTOMER RELATIONSHIPS

It's not enough to have historical records, however. Kubiak says that the next-greatest challenge when trying to derive business value from text analytics is that "records need to be stored with the time stamp or the customer identifier for use later in merging the unstructured data with structured data. Of course, in sources like social media chatter, there is no customer identifier for you to work with. In that case, the only option is to look at the time series, look at the date when certain messages are popping up, and try to align that information with changes in the business or changes in marketing."

"For instance," Kubiak says "When we introduce major changes to our marketing campaign, we want to know how customers react. To do that, we can look at the social media chatter and see what the impressions are, but where we find the most value is when we are able to combine the results of the text analytics with structured data. Potentially, we can use the topics customers are discussing with us as early indicators for shifts in customer preferences and sentiments. We are also trying to improve our customer satisfaction by listening to the voice of the customer."

Although the field of text analytics offers many potential benefits for creating deeper relationships with customers, Kubiak warns that "it's a new and sexy field, but it's not really glamorous. Whoever wants to venture into this area should be patient and ready to spend time preprocessing the data—on things like text parsing, tokenizing. This is a time-consuming and painstaking process but it may have great impact on the results of the analysis."

"They also need to understand the other aspects of preprocessing, like regular expressions, removing special characters, removing some phrases and punctuation from the text," he adds. "That takes a lot of time. Software can do some of it for you, but to become better and better, it's good to have a better understanding of how these techniques work and the cautions and benefits of some of these approaches." To achieve that better understanding, Kubiak suggests, "Text analysts should try to spend time learning what has already been done in academia and look for business applications for these techniques."

“

We can use the topics customers are discussing with us as early indicators for shifts in customer preferences and sentiments.

”

STARTING SMALL CAN BE THE KEY TO SUCCESS WITH TEXT ANALYTICS



**THOMAS
KUNJAPPU**

Senior Product Manager,
Twitter

Thomas Kunjappu loves working on novel products that use data to solve problems. He is currently working on corporate tools @ Twitter, where he imagines a more productive workplace enabled by software. He graduated from Williams College and lives in the San Francisco Bay area.



Download the full e-book: [Text Analytics](#)

Businesses should have clear goals in mind when they are considering text analytics technology, according to Thomas Kunjappu. Rather than getting drawn in by the hype cycle that surrounds text analytics tools, it's essential that business leaders have clear success metrics in place so that, prior to any engagement or initiative, they know what the end state or desired outcome should be.

It's also important to be organizationally ready for text analytics and to know what you're getting into when you're talking about this type of project, he believes. This includes questions of talent resources, industry-specific skill sets, and linguistic approaches as well as software requirements.

In terms of business goals, simply aiming to discover and understand sentiment are not ends in themselves. "We need to go one step further and get to the insights. It's better to start with a smaller area of focus, and then widen your scope," advises Kunjappu. "Then, you'll start to build momentum for the technology and its uses within the company, and you will be confident of the benefits of using the technology."

Kunjappu once worked with a telecommunications client that began with a broad imperative: to know what people were saying about the company and to acquire the necessary software to achieve that one goal. This approach met with mixed results, he says.

“ It's better to start with a smaller area of focus, and then widen your scope. ”

KEY LESSONS

- 1 When considering a text analytics initiative, it's essential that you define clear, focused business goals.
- 2 Starting small can be a smart way to demonstrate the value of text analytics and build momentum for future endeavors within the company.



STARTING SMALL CAN BE THE KEY TO SUCCESS WITH TEXT ANALYTICS

“They came up with their ontology and hundreds of topics in a hierarchical format that the central team thought was relevant and would provide something useful for the rest of the organization. There were lots of pretty reports, but the decision makers and key stakeholders didn’t think that they answered any of the questions that were top of mind for them.” Unfortunately, when those stakeholders voiced the questions that they felt were critical to success, the central team wasn’t ready to handle them.

So, the team adjusted its approach, picking one unit within the company and solving a problem for it. This unit had a specific question about customer churn involving a particular region and product set. The team captured the stakeholders’ hypotheses about why churn was happening, then trained the staff to put in the linguistic rules to accurately build up topics for analysis. Because it had a defined goal that a clear business question facilitated, the organization was able to achieve the results it was looking for.

This development completely changed the conversation about text analytics within the company, Kunjappu says. “The business stakeholders were finding value in what they were doing, and they wanted to invest more in that team. Word got around, and other units were clamoring to be next.” They then performed different analyses for other types of units. “The really interesting insights come when you marry text analytics with quantitative data, operational data, financial data, and—in this case—network data because it was a telecommunications company,” Kunjappu observes. The firm was able to combine those varied data sets in increasingly sophisticated ways, leading to far more innovative insights than were originally possible.

Kunjappu notes that building a global ontology on customer review data is useful for describing what is happening but it often leads to obvious answers that leave the business thinking, “Tell me something I didn’t know.” Talk to stakeholders early to understand their and hypotheses, they use text analytics to prove or disprove them. In this way, you increase the engagement level and the importance of the project immensely.

Businesses can reap bountiful rewards from text analytics initiatives but only when they approach both the strategy and the technology that supports it in a careful, considered manner. By defining clear business goals and accurately identifying the resources needed to achieve them, it is possible to gain such valuable insight. For companies new to the process, starting with a small, focused initiative can often be the best approach.

“

Really interesting insights come when you marry text analytics with quantitative data, operational data, financial data, and—in this case—network data.

”

BETTER THAN YOU WERE



META S. BROWN

President,
A4A Brown, Inc.

Meta Brown's work helps organizations build analytics competence and improve communication between technical and nontechnical professionals. She is author of *Data Mining for Dummies*. You can find links for many articles on analytics and communication on her website, metabrown.com.



Download the full e-book: [Text Analytics](#)

For Meta Brown, the definition of text analytics is simple. "It means converting text into some simpler, more conventional data structure—period," she says. The difference between that and all the other definitions she hears is that the others start there, but then plug in additional layers of meaning, which she thinks really just refer to conventional analytics.

Her guiding principle when it comes text analytics is similarly straightforward: "You will not make any money on text analytics if you do not start with a plan to do that." With that in mind, she offers three pieces of advice:

- **Present a convincing business case.** Identify a case for text analytics that solves a business problem and saves money. Executives often do not want to hear a business case built on revenue increases because they do not believe it will happen, Brown says. They prefer cost savings. Perhaps they have been burned with false promises based on past technology investments and process improvements, she suggests. "Revenue is not the strongest way to make the case that justifies someone agreeing to take on the costs that go into your plan," she adds.
- **Address a specific problem.** Look for a small payoff initially. Some retailers that see a lot of customer service inquiries still deal with those inquiries manually, which creates labor costs, potential opportunity costs, and possible customer satisfaction issues. Instead, she suggests, "We could use text analytics to direct those inquiries, sort them, and deal with them more quickly than through the process we have now." That would result in cost savings and resolve a simple problem, in turn encouraging the organization to wade deeper into text analytics. "Each tiny success is worth something," Brown says.

“ You will not make any money on text analytics if you do not start with a plan to do that. ”

KEY LESSONS

- 1 Executives tend to prefer the cost-savings case to the revenue-generating one.
- 2 Text analytics can be core to a business—but it is no panacea.

BETTER THAN YOU WERE

- **Follow through.** Say that you have proposed investing in text analytics to log and categorize open-ended, written Web survey responses. Do not deviate from your plan. Make sure that any required employee training is completed. Then, use the automated tools. Brown is stunned at how often software subscriptions lapse because companies never get around to using them. Instituting a new system need not require replacing your current processes—at least not immediately—but ultimately it should be out with the old and in with the new. “It’s not enough to buy text analytics software,” Brown says, “you’ve got to learn it, use it, and use the results to guide the business. No change, no benefits.”

Brown has seen first hand how badly her advice is needed. For example, she spent her early days as a trainer, teaching survey research and data analytics to mid-career professionals. In that training, she always asked if their businesses conducted surveys that used open-ended written responses. If they did, she asked which companies sent the responses out to be coded, which merely glance at them, and which ignore them. “In every single group, multiple people admitted that they did nothing with the text,” she says. “They were spending money to collect data that they weren’t using.” That is a waste of resources and customer goodwill, she states.

In contrast, approaching text analytics as a fail-safe answer would also be a mistake. Businesses sometimes abandon text analytics because of that misunderstanding. In one case, Brown notes, an organization uncovered a single mistake when an automatic categorization turned a positive term into a negative one. After that, the company rejected text analytics altogether.

Text analytics can become a core part of your business process, but make peace with its imperfections, she says. “You are not going to be perfect—nothing else in life is perfect. The objective is to be better than you were before.”

“

In every single group, multiple people admitted that they did nothing with the text. They were spending money to collect data that they weren't using.

”

WHERE DO YOUR HIGHEST-VALUE CUSTOMER ENGAGEMENTS OCCUR?



**BILL
SHELDON**

Chief Solutions Officer,
Angoss

Chief solutions officer for Angoss' predictive analytics professional services division, Bill Sheldon has more than 20 years of experience in sales and marketing effectiveness. Prior to joining Angoss, Bill was vice president of Marketing Solutions for Paradyz Analytics, a multichannel marketing services agency, and chief operating officer for CFM Partners, a leading provider of Software as a Service compliance risk-management solutions.



Website



Download the full e-book: [Text Analytics](#)

Bill Sheldon's view on the business value of text analytics is shaped by his role as a business-to-business (B2B) provider of predictive analytics to the financial services industry. "We work with a lot of asset management companies, companies that sell mutual funds through a network of financial advisers," Sheldon explains.

Mutual fund companies have internal salespeople who call on financial advisers to promote different funds that achieve different financial objectives. Some companies use an analytics model to tell them which advisers are most likely to purchase and which products they will buy and to identify upsell and cross-sell opportunities. "When we engage with customers," says Sheldon, "we look specifically at their challenges, which are mostly in sales and marketing." The problem you need to solve points to the kind of data you need to solve it, and text is a valuable part of that data set. But it is just one of many data sources. "Text is not typically stand-alone data," says Sheldon. "It's an important companion to other structured business data."

“Text is not typically stand-alone data. It is an important companion to other structured business data.”

KEY LESSONS

- 1 The problem you need to solve points to the kind of data you need to solve it, and text is a valuable part of that data set.
- 2 Increasingly, companies are recording their customer conversations and using speech-to-text tools to convert them for text analysis.



WHERE DO YOUR HIGHEST-VALUE CUSTOMER ENGAGEMENTS OCCUR?

Text analytics is especially valuable in discovering, issues, opportunities, and trends. For instance, you can look at customer service notes to see who's complaining about what. Or you can look at customer relationship management (CRM) activity that salespeople enter after sales conversations and emails between salespeople and their customers. "When we look at email, we look at sales and customer emails separately," says Sheldon. By looking at emails in this way, you can remove the sales bias and get a better sense of the true customer voice. Text analysis can also serve as a broader market-sensing tool to reveal trends, such as a growing interest in certain kinds of investment products.

"These kinds of analyses can be brought back to business performance, and then we can look at this in several different dimensions," says Sheldon. For example, from a product dimension, you can look across all those interactions and see how people are feeling about specific products. Or you can look at sentiment in the sales dimension. This could mean looking at all the conversations of one salesperson to see which phrases that person uses and perhaps do a comparative analysis of different salespeople and tease out differences in what they are saying that might give insight into differences in their performance. Sheldon says, "Text analytics used in this way becomes a valuable sales coaching tool."

Many companies, particularly in the B2B segment, spend an extraordinary amount of money putting information such as meeting notes into their CRM systems, but they spend little money extracting insights from those data. Mostly, they use call meeting notes as memory refreshers before their next engagement, but Sheldon points out, "Those meeting notes contain a lot of valuable business insight."

Many companies don't know the value of their text, or if they do, they don't know which text they should analyze. "When companies are looking to extract value from their text data," says Sheldon, "they should begin by looking at where the highest-value customer engagements occur."

Increasingly, companies are recording their customer conversations and using speech-to-text tools to convert them for text analysis. This practice is especially important in B2B businesses like financial services, where highly compensated, skilled salespeople speak with highly skilled investment advisers. "These are high-value conversations that can give you a lot of insight into all the dimensions of that engagement," says Sheldon. "In text analytics, the greater the volume of text you have to work with, the more insight you can derive from it."

“

When companies are looking to extract value from their text data, they should begin by looking at where the highest-value customer engagements occur.

”

THE VOICE OF THE CUSTOMER



**KAENAN
HERTZ**

Executive Director,
Ernst & Young

Kaenan Hertz specializes in helping clients use strategic information and analytics to drive innovation in products and services. He has extensive experience helping companies innovate their customer experience strategy; enhance offerings through digital, traditional, contact center, and third-party channels; and understand the implications of changing market and competitive dynamics on customer strategies. Kaenan has held leading roles in marketing and decision management for top insurance, banking, and finance companies, including USAA, Citibank, and Sallie Mae.



Download the full e-book: [Text Analytics](#)

Before taking his present job at Ernst & Young, Kaenan Hertz was a market intelligence executive at a top-tier financial services company. There, he once headed up a “voice of the customer” project that identified more than 100 points of entry for customers to offer feedback, including letters to the chief executive officer, conversations with customer service representatives, emails, surveys, and touch points across all departments and product categories. Hertz’s project aimed to generate a 360-degree customer view by building the infrastructure to gather, synthesize, and analyze customer feedback into a central database. “You can then start to use that database to drive insights,” he says.

Some of those insights were unexpected. Hertz’s former company offered several automobile insurance products, for example. After analyzing customer feedback, the company was surprised to discover that some of its insurance feature sets were not important to customers. For Hertz, the moral of that story is that you ultimately cannot do text analytics in a vacuum. “It is a collaborative partnership among many different departments in any given company,” he says.

“ They never drive value . . . because they haven't articulated the business problem they are trying to solve. ”

KEY LESSONS

- 1 Text analytics will ultimately require broad interdepartmental collaboration.
- 2 It would be wise to start with a small pilot project and grow your capabilities from there.



THE VOICE OF THE CUSTOMER

In that spirit, Hertz offers some advice for getting your text analytics journey underway:

- **Start with a real business problem.** Only when you know the issue you want to resolve can you find the appropriate tools to answer it, Hertz states. At Ernst & Young, he has seen companies become so enamored by new tools and technologies that they forget to connect them to any real business objectives. “So they never drive value,” Hertz observes. He is personally aware, he says, of several companies that are now limping through big, unstructured text data projects. Why? “Because they haven’t articulated the business problem they are trying to solve,” he states.
- **Start small, and grow from there.** He calls this approach the *crawl, walk, run methodology*. Although ultimately, text analytics cannot be done in isolation, you can begin modestly. An example might be collecting open-ended feedback and distilling it into root causes for a customer’s most recent transaction. That would answer a single problem using one well-defined source of information, he says. You would not need to merge data across multiple systems and different departments. Small successes like that can generate corporate enthusiasm to keep pushing forward. Conversely, he warns, “If you try to boil the ocean, you will never succeed.”
- **If you collect data, use them.** If you solicit the customer’s input, you need to process it and act on it, Hertz states. That is what customers expect. After all, he says, performing text analytics is really about listening to the voice of the customer. “Don’t just collect information and do nothing with it,” he concludes.

“

If you try to boil the ocean, you will never succeed.

”

AVOIDING CATASTROPHE



**SHREE
DANDEKAR**

Executive Director, Product Management, Predictive Analytics, Dell Inc.

Shree Dandekar has been at Dell for the past 15 years in a number of roles, from software design, product development, and enterprise marketing to technology strategy. Currently, he is the executive director of product management and strategy, responsible for developing and driving the strategy for Dell's business intelligence and analytics solutions.

 |  | 
Twitter | Website | Blog



Download the full e-book: [Text Analytics](#)

There is a common misconception that text analytics is useful only in the social media space, says Shree Dandekar, executive director for data analytics at Dell. That is false, he asserts.

“Text analytics is almost like a foundational block for creating a solution that has any kind of semistructured or unstructured data,” Dandekar states. “Nobody can deliver true solutions without it.”

The technical and business goals for much of Dandekar’s work at Dell are tailored around social media, but he looks well beyond that when offering his advice for deriving business value from text analytics. His suggestions:

- **Listen in.** Listening to customers means more than monitoring Twitter, Facebook, and LinkedIn. Those platforms contain only 40 percent to 50 percent of the data set, Dandekar says. Much of the conversation takes place elsewhere—on online community sites, chat sessions, customer support logs, and other vehicles. “Where the serious conversations happen is in some of the back-end communities,” he states. “That is critical when you are talking about social media analytics.”

“ Where the serious conversations happen is in some of the back-end communities. ”

KEY LESSONS

- 1 Listening to the pulse of the customer base requires more than social media monitoring.
- 2 You can glean critical business insights from any semistructured or unstructured data set. A hybrid model is the ultimate goal.



AVOIDING CATASTROPHE

- **Monitor trends.** Collect, record, and analyze textual data. “When you have your data set, you have to spend a lot of time modeling the data,” Dandekar states. Tools are available, for instance, that can assemble word clouds to help sniff out top trends and topics. Such tools are helpful, but Dandekar prefers building taxonomies to generate a hierarchy of relevancies. “You can filter your entire data set against a predefined taxonomy to help you understand where all the conversations are taking place,” he says. “That’s the first step in creating a high-level data model.”
- **Contextualize.** Unstructured data require context to produce actionable insights, according to Dandekar. Natural language processing–powered sentiment analysis, for example, can identify when conversations about your company or its products subtly take a negative turn. “Without the business context,” he says, “the data are not going to make sense.”

Dandekar describes an instance where context helped stave off a consumer revolt. At the time, Dell was set to release a notebook PC geared toward developers; it had announced that an open source operating system would be installed on the computer. Early social media feedback was glowing. Late in the process, however, a pricing manager failed to take into account the open source operating system when announcing the machine’s price. Online conversations began to sour, a fact that was revealed only because Dell was monitoring conversations with the aid of text analytics.

Dell had developed taxonomy for the machine that allowed it to drill down by topic, Dandekar says. When it came to the subject of price, the notebook’s Social Net Advocacy (SNA) sentiment scores were tanking. Realizing what happened, Dell took action. Within 24 hours, the price was reset and apologies were issued by email, Twitter, and blogs. The SNA scores rose.

For Dandekar, that incident’s implications extend well beyond social media. Critical business insights can be gleaned from any unstructured business data—not just social media, but also customer relationship management data, enterprise resource planning data, and transactional analytics. In fact, he says, the boundaries around customer data sets are collapsing.

Creating a hybrid data analytics model, then, is the ultimate goal, but for now, Dandekar observes, that goal remains elusive. It is the subject of a much bigger philosophical discussion that many businesses—Dell included—must have soon.

“

When you have your data set, you have to spend a lot of time in modeling the data.

”

THE VALUE OF TEXT ANALYTICS COMES FROM BUSINESS CONTEXT



**RAMKUMAR
RAVICHANDRAN**

Director, Analytics & A/B
Testing, Visa Inc.

Ramkumar Ravichandran is the director of analytics and A/B testing in the Digital Product organization at Visa. His core responsibility is to enable a data-driven decision-making framework for Visa's leadership and stakeholders. Ram is passionate about the value of data and insights, and he loves sharing what he learns at conferences, in blogs on LinkedIn, and through Slideshare. He is also writing a book titled *Practical Predictive Analytics*, an overview of the analytical techniques that help deliver value to organizations.



Download the full e-book: [Text Analytics](#)

Ramkumar Ravichandran's great revelation about the power of text analytics came when he joined a company in which research, the purview of text analysis, and business analytics fell under one manager. "That was the first time I realized the power of text analysis in complementing what we get from business analytics," says Ravichandran. Analytics is a numbers story that comes out of an analyst's head, but the reasons *behind* that story are based on what customers say. "Only when we put these two together can we get the complete picture of customer actions that show up in the analytics as numbers." Text analytics is not going to answer all business questions; it is a complementary technique that completes a picture, not a solution by itself.

A good example of how text analytics fits into a larger business process is how it serves a product development life cycle, which encompasses many business activities and types of analyses. It might begin with an envision stage, where people decide what they are doing and why they are doing it. There would be a review of business reports that provide metrics related to the business activity under consideration. There would be analysis of text from a variety of sources, such as social media, call center feeds, surveys, and reviews. This analysis would provide deeper insight into justifications of the product's development from the customer's perspective.

“It's only when you look at other sources of information and insights that can you turn text analysis into a business recommendation that will benefit the company.”



KEY LESSONS

- 1 Analytics is a numbers story that comes out of an analyst's head. You must look to what the customers say to find the reasons behind that story.
- 2 You must be able to translate the analysis back into actionable recommendations.

THE VALUE OF TEXT ANALYTICS COMES FROM BUSINESS CONTEXT

Based on this analysis, there might be several recommendations for possible approaches to the problem, and then A/B testing to narrow the choices down to one solution. The final decision would be based on analytical and customer data that show why this is the optimum approach. "It's only when you look at other sources of information and insights that can you turn text analysis into a business recommendation that will benefit the company," says Ravichandran.

To make text analytics really impactful, you must:

- **Consider your business model.** For instance, a retail business will rely on different kinds of text data than a business-to-business organization.
- **Understand the business need.** Is the analysis related to product development or user experience management? Is it going to improve an operational process?
- **Understand the questions you are trying to answer.** This involves considering actions you might take based on answers the analysis provides. You must also have a good framework for setting expectations. Text analysis does not answer all business questions, and a lot goes into preparing clean data for analysis. Finally, you must be able to translate the analysis back into actionable recommendations.

Text analysis can reveal unexpected results. Ravichandran cites an example in which an analyst was analyzing product reviews from customers. He decided to look at just five-star reviews to identify positive experience drivers. In the course of looking for what made customers happiest, he discovered that although these customers really liked the product, there was a pattern of comments revealing a problem with shipping. "This research, which was designed to find positive experience drivers, identified a problem that could easily be fixed through an operational process change," says Ravichandran.

Ravichandran advises that when doing business analysis, it is important to connect analytical dots from different places, then complement them with text analysis to answer the question, "Why."

**Participation in this e-book is purely on personal basis and does not represent Visa in any form or matter. The advice in this article is based on learning from work across industries and firms. Care has been taken to ensure that no proprietary or work-related information of any firm is used in this material.*

“

This research, which was designed to find positive experience drivers, identified a problem that could easily be fixed through an operational process change.

”

TEXT ANALYTICS OFFERS POWERFUL TECHNIQUES FOR REVEALING HIDDEN INSIGHT



**SETH
EARLEY**

CEO,

Earley Information Science

Seth Earley is CEO of Earley Information Science, an information management strategy consulting firm and the editor of data analytics for *IT Professional Magazine*. He also serves on the editorial board of *Applied Marketing Analytics*. His interests include knowledge strategy, data and information architecture, search-based applications, and information findability solutions. Seth conducts workshops for senior leadership on aligning information management strategy with measurable business outcomes and develops information governance programs for clients in multiple industries.



Twitter



Website



Blog



Download the full e-book: [Text Analytics](#)

There are myriad use cases for text analytics in verticals ranging from retail and customer service, to life sciences, telecommunications and far beyond, according to Seth Earley. Many techniques and tools are available that allow practitioners to make sense of large amounts of content – such as email messages, customer service tickets or processes large numbers of research studies to identify and understand scientific trends. The key to realizing value is in applying structure to unstructured data and knowing what questions your company seeks to answer in the data and content.

If you are looking at a body of content that is well written or vetted, it is easier to find the structure. A news article is well organized (compared with a typical email message for example) and usually is about a specific thing or event. One way you can draw insight from it is by performing entity extraction. Entity extraction works best on well-formed content but can also work on less structured information like social media posts. “By creating a list of terms that are of interest – such as your products or those of a competitor (also known as a *controlled vocabulary*),” says Earley, “you can group comments according to those term occurrences and focus on the things that people say about your product or those of your competitor. You can also identify whether comments are positive or negative using sentiment analysis – another text analytics approach.”

“ You can parse the information according to a process or problem, and then correlate that data with internal processes. ”

KEY LESSONS

- 1 Text analytics involves finding structure in unstructured data so that you may properly analyze them.
- 2 There are many powerful techniques and tools that practitioners can use to uncover valuable business insight from their unstructured data.



TEXT ANALYTICS OFFERS POWERFUL TECHNIQUES FOR REVEALING HIDDEN INSIGHT

You can process a wide range of content, including blog posts, discussion forums, trouble tickets, customer service calls, and transcripts of call center conversations. “You can parse the information according to a process or problem, and then correlate that data with internal processes,” Earley explains. “If a user posts a complaint and describes how your product arrived late, was damaged in shipment, that they then did not hear back from customer service, then called and waited on hold for 20 minutes to reach an agent, and it took 4 weeks to get their refund, there are several internal process issues which that one piece of feedback identified. Pulling these kinds of insights from dozens or even a hundred email messages is one thing. Trying to process thousands or tens of thousands requires text analytics to identify themes and process issues. This is how voice of the customer data can be mined from unstructured text and used to identify problems in various departments.”

There are specific business problems, even seemingly insurmountable ones, to which text analytics can be effectively applied, notes Early. A field service organization he once worked with was taking days to service equipment because the technical documentation they used was difficult to find and it was located in several different places. The organization had attempted to solve this problem several times before with no luck. It wasn’t until they took a structured approach to analyzing the text and the content that they were able to reduce the amount of time that their engineers were looking for information by 50 percent.

The field service technicians were spending 40 percent of their time—approximately 16 hours a week—searching for information that they could not find. After categorizing the data so they could retrieve the specific content they needed, they were able to save eight hours per week per technician. With 3,000 service engineers across the organization, that amounted to \$50 million per year in savings.

In another example, Earley once discovered incredible business value derived from text analytics during a project for a life sciences firm that was trying to locate research about a particular drug. “They didn’t have visibility into it because it was deep on the 30th page of search results,” Earley says. “But by doing some entity extraction, they were able to find underutilized content that allowed them to pursue some new research avenues for some medications that were going off patent.” Through applying auto-categorization techniques, they were able to mine the organization’s deeper content and reveal vital, actionable information.

Text analytics is a powerful tool for unlocking hidden business value and solving complex business problems across a range of industries. By applying a set of best practices to a clearly defined business problem or question, you can achieve dramatic outcomes to move your business forward.

“

By doing some entity extraction, they were able to find underutilized content that allowed them to pursue some new research avenues.

”



**SUDIP
CHAKRABORTY**

President,
Xypress LLC

Sudip Chakraborty founded Xypress to help companies grow profits by applying sophisticated analytics to big data. He and his team have helped companies uncover insights from structured and unstructured data in the financial services, technology, information services, and pharmaceutical industries. Sudip holds two patents in the areas of approximate text matching, text analytics, and association discovery: [System and process for discovering relationships between entities based on common areas of interest](#) and [System and method for searching and matching databases](#).



Twitter



Website



Blog



Download the full e-book: [Text Analytics](#)

As an expert in applying sophisticated analytics in business-to-business (B2B) settings, Sudip Chakraborty, president of Xypress LLC, offers four recommendations for getting the most out of text analytics:

- **Develop a set of hypotheses regarding how you can solve a business problem.**
- **Apply text analytics and review and refine your original set of hypotheses.**
- **Leverage humans to evaluate, refine, guide and apply – machines are not sufficient.**
- **Deploy a cross-functional team of highly trained professionals with skills in business problem solving, text analytics and structured data analytics.**

Develop a set of hypotheses. As a first step, generate three to five hypotheses about how text analytics might help your business to solve one or a set of business problems. The challenge, he states, is that there is no end to the possibilities: “Text analytics consists of many different techniques including sentence segmentation, parts of speech tagging, named entity recognition, sentiment analysis among others - you can, therefore, spend days, weeks, even months, applying text analytics to unstructured data without realizing business value from your efforts.” He also notes that the analysts who understand and can solve text analytics problems often have different educational backgrounds and experience from those who understand and solve business problems. It’s up to business decision makers to set direction and time limits. “Analytics professionals can spend a lot of time looking for things,” Chakraborty observes. “Unless they get direction, their findings are unlikely to be of business value.”

“ We actually needed business analysts to evaluate the quality of the results and to ensure that they would be valuable to our client's customers. ”

KEY LESSONS

- 1 **Develop a set of three to five hypotheses for text analytics, then be prepared to iterate.**
- 2 **The real business value of text analytics comes from analyzing both structured and unstructured data.**



FIRM, NOT FROZEN

Iterate. Your theories should be firm, not frozen. In particular, if you have little prior experience, you might be far off target at first. Redefine your hypotheses periodically. Chakraborty illustrates his point with a story about how he once helped a \$50 billion tech company explore the wearables market. The client directed its own considerable resources to interviews with experts and focus groups. Chakraborty's team, meanwhile, performed social media analytics to unearth trends in wearable technology. While performing this analysis, the team uncovered a problem no expert had yet spotted: competitors' wearables caused skin irritation for some customers. However, this was not part of the original hypotheses. Chakraborty's team worked with the client closely to refine the original set of hypotheses; they decided to apply a broader set of text analytics techniques to dig deeper into the skin irritation issue. The client, therefore, had an opportunity to reassess its device before it ever hit shelves.

Software cannot do it all. Human assessment is crucial. Again, Chakraborty speaks from experience. He holds a patent on a sophisticated B2B search engine designed for a global wholesale website that performs product searches through association discovery and uncovers purchase trends. During that project, he felt compelled to engage business analysts beyond the original team of data scientists and software developers. "Just looking at the data that our text analytics threw up was not sufficient," Chakraborty says. "We actually needed business analysts to evaluate the quality of the results and to ensure that they would be valuable to our client's customers."

Deploy a cross-functional team with complementary skills. People tend to approach data analytics as an either/or proposition, he observes—it's all about structured or unstructured data. That approach risks sacrificing the real business value of text analytics, which comes from analyzing both structured and unstructured data, Chakraborty states. "In my experience I see that organizations tend to start with one knowing that they will get to the other," Chakraborty says. "It's just that it is often challenging to do both together, and they don't want to take on more than they can handle." Humans, he adds, remain indispensable. "Machines are increasingly able to perform significantly more intelligent tasks, but be prepared for humans to assess, refine, and guide the execution of your text analytics program," Chakraborty concludes. "That is the way that you will actually maximize the business value from your efforts."

“

Machines are increasingly able to perform significantly more intelligent tasks, but be prepared for humans to assess, refine, and guide the execution of your text analytics program.

”

MEANINGFUL ANALYSIS DEPENDS ON CHOOSING THE RIGHT DATA



**DANIEL
ANGUS**

Lecturer in Computational
Social Science,
The University of
Queensland

Dr. Daniel Angus is a computational social scientist who specializes in the design, development, and application of advanced computational methods for the study of communicative practices. He is a co-inventor of the Discursis software technology, which has been used to reveal the dynamics of topical exchange in health, education, media, and political conversational contexts.



Twitter | Website



Download the full e-book: [Text Analytics](#)

Dr. Daniel Angus, who began his career as a computer scientist, came to text analytics in an interesting way. As a postdoctoral research fellow, he became involved in researching how animals navigate. Evidence suggests that many aspects of physical navigation are governed by the same part of the brain that processes language, the hippocampus. That revelation was the basis for a whole new approach to language and text processing. “We began thinking about processing communication data by looking at it through the lens of navigation,” explains Angus. This approach has turned out to be especially useful in analyzing conversational texts.

This kind of text analysis is used in many ways. For example, you can learn a lot that is relevant to your business by listening to how people talk about your brand. Text analytics allows you to isolate your brand name as a concept, then using a highly statistical approach to see what concepts are associated with it. Through co-locating concepts and words, you can quickly measure the kinds of words people are using in close proximity to your brand. “This not only provides insight into what people say about your brand, but it might reveal deeper insights. It could be a first step in exposing why people are thinking it,” explains Angus. Reasons “why” can have deep and not-so-obvious roots that may be different for different population segments.

“ This not only provides insight into what people think about your brand, but it can go a lot deeper. It can expose why people are thinking it. ”

KEY LESSONS

- 1 The real cost of text analysis is not in the analytical tools but in cleaning the data.
- 2 Analyzing unfiltered text frees the analysis from human biases that might creep into the analytical process.



MEANINGFUL ANALYSIS DEPENDS ON CHOOSING THE RIGHT DATA

Text analysis offers a different kind of insight than that provided by structured data analysis. “When we start talking about the difference between quantitative and qualitative research, this kind of text analytics blurs the distinction because it allows a very rich qualitative insight,” says Angus. It can have broad application in understanding deep sentiments and motivations behind population segments as they relate to complex issues, including public affairs, political issues, and many kinds of business questions.

When considering text analytics, Angus stresses the importance of knowing your goal. Only when you ask yourself, “What is a really interesting question or goal for me within my organization?” are you able to decide whether text analysis is appropriate for this kind of question; which are the right analytical tools to use in approaching that problem; and, if it involves text analysis, what text should be analyzed.

Angus points out that the question of what text to analyze is important. In instances where you are analyzing what people say, it’s best to analyze complete transcripts rather than notes because notes invariably introduce the bias of the note taker, and may miss seemingly mundane, but potentially important details. However, says Angus, “This is where the big cost question comes in.” The real cost of text analysis is not in the analytical tools but in cleaning the data. There is the challenge of accurate speech-to-text conversion, formatting issues associated with online data, and print text from news services, which comes in a variety of formats. “People often underestimate the time and cost of actually collecting and cleansing the data.”

For meaningful analysis, the text you chose is critical. Analyzing unfiltered text frees the analysis from human biases that might creep into the analytical process. “It’s not about removing human judgment,” explains Angus. “It’s about revealing true, genuine patterns, and then using human judgment to interpret those findings.”

“

People often underestimate the time and cost of actually collecting and cleansing the data.

”

ROBUST TEXT ANALYTICS RESULTS REQUIRE A HUMAN ELEMENT



**HELEN
CLEGG**

Social Analytics Manager,
A.T. Kearney

As social analytics manager at A.T. Kearney, Helen Clegg uses her library science background and expertise in taxonomy development to add value to the text analytics capabilities of the company's Data Science Team. She is a regular presenter at Taxonomy Boot Camp and Text Analytics World and a guest lecturer at the Grenoble Graduate School of Management and the London School of Business and Finance. Helen is an organic gardener and studying for her Permaculture Design Certificate.



Twitter | Website



Download the full e-book: [Text Analytics](#)

Helen Clegg, social analytics manager at A.T. Kearney, had no idea when she chose to pursue a degree in library science that she would eventually be using that knowledge to improve text analytics results. That is exactly what happened, however, when she began working with text analytics at A.T. Kearney, and now she has a unique view of how to gain the greatest business value from text analytics efforts.

“I think companies are struggling to get value from text analytics for a number of reasons,” she says. “First of all, it’s not easy to translate a text analytics concept into business terms. It’s really unfamiliar territory, and if you think about it, the language that you use when you talk about text analytics is unfamiliar language—terms like *semantics*, *lexical chains*, *algorithms*. When you start talking about these concepts and use these terms, people don’t understand it. It’s unfamiliar, and so they don’t know how to get business value from it.”

It’s doesn’t have to be overwhelming, however. Clegg suggests, “I would recommend going for a project-oriented approach. You pick a project or a use case, if you will. You start small, and you work out what the business problem is. What is the problem that the project team is trying to answer?”

“ I would recommend to go for a project-oriented approach You start small, and you work out what the business problem is. ”

KEY LESSONS

- 1 A good approach is to take one small element of a larger project and work toward finding the right solution for that element, expanding text analytics efforts over time.
- 2 Machines don’t have common sense. They don’t understand subtly and nuances. For this reason, the human element is still one of the most important elements of text analytics.



ROBUST TEXT ANALYTICS RESULTS REQUIRE A HUMAN ELEMENT

“By doing that, it makes the concept of text analytics much more tangible and demonstrates, if you’re successful, the speed and just how quickly text analytics as an application can help solve that problem,” she says. Clegg also says that this type of approach makes it easier to determine which text analytics tools to use. “This is important,” she says, “because there are lots of tools in the marketplace. We started with a couple of open source tools, and now we’ve added a couple of licensed products, so we have a mix. It’s only by taking this approach that it’s allowed us to work out the tool set that we need.”

Although Clegg finds value in analytics tools, she warns that for truly robust results, text analytics must also have a human element. “Don’t discard the importance of the human element. There are many proponents out there who say use a machine-only approach. They let the big data tools and the algorithms do everything, but we’ve found that our results are even better and they’re more robust if we leverage human intelligence during the process. We’re leveraging classical library science skills and that’s where I fit in. I have the library science qualification, and that skill is really useful for designing a custom taxonomy for whatever project or use case we’re working on.”

Clegg points to an example that was used as an illustrative exercise for a project team. “This particular team wanted to know what was top of mind for a particular group of consumers when buying a specific brand of car,” she explains. “This required mining a vast and complex array of social media posts. Leveraging human intelligence, we developed a custom automotive taxonomy based purely on what the project wanted to find out. Once we had the custom taxonomy worked out, we were able to mine these social posts with this taxonomy. That enabled us to gain rapid insight into consumers’ buying behavior.”

One last bit of advice from Clegg, “I would also say don’t get hung up on sentiment analysis. There are quite a number of sophisticated sentiment engines on the market but what we would say from the work we’ve been doing is that sentiment analysis doesn’t always provide the answer to the business problem. Saying that something is negative or positive doesn’t really give you deep insight,” she says. “That’s why we have developed more of a contextual approach where we leverage human intelligence around the taxonomy and combine it with text analytics on the machine side.”

“
Sentiment analysis doesn't always provide the answer to the business problem. Saying that something is negative or positive doesn't really give you deep insight.”

AVOID DRAWING INCORRECT CONCLUSIONS



**JEFF
CATLIN**
CEO,
Lexalytics, Inc.

Jeff Catlin has more than 20 years of experience in search, classification, and text analytics products and services and has held technical, managerial, and senior management positions in such companies as Thomson Financial and Sovereign Hill Software. Prior to forming Lexalytics, Jeff was general manager of LightSpeed Software, where he was responsible for sales, marketing, and development efforts for the Knowledge Appliance and iFocus products. He holds a bachelor's degree in electrical engineering from the University of Massachusetts Amherst.

 |  | 
Twitter | Website | Blog



Download the full e-book: [Text Analytics](#)

The value of text analytics is not in the *what* but in the *why*. “You can dig out facts, but if you don’t go deep enough, the facts won’t tell you much,” says Jeff Catlin, whose business specializes in analyzing various kinds of social content. In fact, you are likely to draw incorrect conclusions if you do not find the deeper context behind what you are seeing in a text analysis.

A good example comes from a project in which a hotel chain wanted to examine customer feedback to see how two Las Vegas hotels—Bally’s and Bellagio—fared with their customers. The project involved applying text analysis to the customer feedback, with particular attention to sentiment about the hotels in general, and specific features like the rooms, the pool, valet service, and location. Such an application is typical of text analytics on social media. An interesting and unexpected finding was that consumer sentiment around Bellagio, a luxury venue, was showing higher negatives than sentiment around Valley’s, a low-cost venue. Valley’s customers seemed to be more satisfied. “We knew the data were accurate,” says Catlin, “because the hotels scored the same on location. They are directly across the street from each other.”

“You can dig out facts, but if you don't go deep enough, the facts won't tell you much.”

KEY LESSONS

- 1 It is quite possible to draw incorrect conclusions if you do not find the deeper context behind what you are seeing in a text analysis.
- 2 You need to use text analytics tools that can understand and correlate different ways of expressing a sentiment. Only in that way is it possible to lift the *why* answers out of the top-level results.



AVOID DRAWING INCORRECT CONCLUSIONS

A first look, at the data suggested that Bellagio had a problem. Here was a cut-rate hotel scoring higher in customer satisfaction than the luxury hotel across the street. A deeper dive into the data, however, revealed that the people complaining about Bellagio were complaining almost exclusively about the price of things at the hotel. Because Bellagio caters to the very high end of the market, they realized that negative ratings based on complaints about price were not a problem for them. “This is a good example of how the data were misleading,” says Catlin. “If you were doing a superficial social media look, those negative sentiments could be worrisome,” but by looking more closely at the data and viewing them in the context of what was actually driving the negative sentiment, you could see an answer to the *why* questions behind the negative ratings. In this case, it became clear that the negatives were not something Bellagio needed to worry about.

“To do this deeper analysis, you need to be able to understand what’s being talked about, even if it is being talked about in seven different ways,” says Catlin. You need to use text analytics tools that can understand and correlate different ways of expressing a sentiment. Only in that way is it possible to lift the *why* answers out of the top-level results.

Catlin says, “Most people who come to text mining think in terms of buckets they understand and problems they know they want to solve. In those cases, it’s usually viewed as a categorization and sentiment problem.” A lot of the more interesting data come out of a “what are people talking about” kind of analysis. You can’t precategorize or build a taxonomy to capture what people are going to be talking about; you have to be able to extract the bits of information that are relevant and hot at that moment. Customers don’t tend to think about text analysis in terms of things they hadn’t expected, but that could harm them. “In my view, that is one of the best uses of this technology. It is totally data driven. You don’t have prebuilt biases about what is going on. The data are just telling you what people are saying.”

“

To do this deeper analysis, you need to be able to understand what's being talked about, even if it is being talked about in seven different ways.

”

What are your customers saying about you?



Receive a complimentary analysis of your text data!

Unstructured data can be overwhelming and difficult to interpret. Our exclusive Mighty Guides FREE offer will help you make sense of your data so that you can make better business decisions!

[Click here to access this offer!](#)

