

Fire Up Your Social Media Strategy with Big Data Analytics

Troy Hiltbrand



Troy Hiltbrand is chief digital officer at Kyäni, and an associate editor of the *Business Intelligence Journal*.
thiltbrand@kyanincorp.com

Abstract

In the age of digital business, the Internet has become a platform for connecting businesses with customers on a much more personal level. Social media act as public forums where customers who are looking for answers can be united with businesses providing solutions. To take advantage of this new form of communication, businesses must develop a social media strategy. With the deluge of information available on social media, analytics is key to enabling businesses to home in on the right users at the right moment for the right engagement.

Introduction

In the era of digital business, the Internet has become less of a delivery mechanism for marketing content and more of an interactive tool that connects businesses with existing and potential customers. One of the key fundamentals of this digital business revolution is the use of social media—a broad term that covers multiple venues where end users actively and collaboratively create and manage information in a public or semi-public setting. Robust social media platforms such as Twitter, Facebook, Pinterest, Instagram, and Tumblr give businesses and customers a whole new way to engage with each other.

What does big data analytics have to do with social media and, more important, what does it mean for the ways businesses interact with their customers? Big data analytics provides the tools to evolve a business's social media role from observer to active and focused participant, getting the right message to the right person at the right time.

The key to success in this new digital business landscape is to have a cohesive and forward-looking strategy for both social media and big data analytics. A company's

social media strategy can and should cover multiple areas, including social media advertising, social media monitoring, and social media engagement. What are the differences among these areas and why are they important?

Big data analytics provides the tools to evolve a business's social media role from observer to active and focused participant, getting the right message to the right person at the right time.

Social media *advertising* is the use of social media to disseminate your advertising message. This can include paid ad placement on social media platforms or the development of product- and company-specific pages that allow your customers to passively consume information or actively share that information with their friends and followers. This is the first step many businesses take when they venture into social media. They use the marketing fundamentals they have been employing for years in a new medium. Customers are spending more of their time in this environment and getting your advertising message in front of them can make a significant impact.

The next step in building your social media strategy takes much more active engagement with customers in understanding who they are and what they are saying. Social media *monitoring* includes capturing and assimilating the voice of your customer as it relates to your business and products. Often enterprises will “listen” to what the community is saying about them and their products and allow those messages to help formulate their other business strategies: product development, market segmentation, and customer relations.

The next generation of business use of social media is to reach out and interact with customers; this is referred to as social media *engagement*. Gartner's description

indicates that it “go[es] beyond monitoring brand mentions by adding [the] capability to create and respond to posts from popular social networks. In advanced tools, the social data is captured and analyzed to determine the incoming post's content, who is submitting the post, the quality of the engagement[,] and additional capabilities for prioritization and workflow.” (Herschel, Chandler, & Sussin, 2014)

If you use social media to engage your customers, you must be able to actively respond in near real time. To do so, your advanced big data analytics must move beyond counting the number of posts about a product or containing a specific hashtag and understand the meaning behind the content. After understanding the meaning of the content, your business can act to further develop the relationship with your customer.

At a small scale, this can be handled manually by a single individual or a small team, but as the number of interactions increases, analytics and automation are needed to help synthesize the incoming traffic and respond quickly. In the same research, Gartner indicated that social media engagement tools will transform how companies do business in the future. This transformation will become mature across the industry within two years. The implementation of advanced analytics to take a social media strategy has the potential to fundamentally alter the future course of your business.

Although engagement is important, author Mark Schaefer points out in *Social Media Explained: Untangling the World's Most Misunderstood Business Trend*, that “engagement is not strategy ... engagement has to be tied to business objectives.” Implementing a solution to have a conversation with customers via social media is not the target; *having a meaningful engagement tied to a business objective is the end goal* (Schaefer, 2014).

Those business objectives can include customer acquisition, new product development, customer service, brand awareness, or investment. There must be a tangible goal and your company must measure its progress as it implements its social media engagement strategy. Otherwise, it will end up spending valuable resources creating

conversations that will get lost within the sea of millions of other conversations that occur every day on these social media platforms.

To advance from social media advertising to social media engagement, your business needs to develop your tactical approach to harvesting social media content, how you will understand the message behind the content and isolate those conversations that are important in achieving your business goals, and how to respond to further develop that relationship so you achieve your objective.

Strategy Step 1: Harvest and Ingest Data

The first step in any social media strategy is to understand what social media communities are available, who the audience is, and how open the social media platform is for data harvesting.

The most popular social media platforms are similar in nature but significantly differ in both how their audience uses them and their openness to allowing automatic harvesting of content.

Facebook

Facebook is the largest social community with over one billion users worldwide—assuring a broad and diverse group. Facebook allows individuals to share photos, videos, and text-based status updates.

The Facebook user community is formed through friendships, which are bi-directional. Both users must opt in to the relationship. Once the relationship is formed, both users have access each other's content.

In addition to allowing individuals to communicate, Facebook lets enterprises create a presence as a gathering place for customers and followers to learn about (and interact with) them. This relationship is formed when an end user “follows” such an enterprise or a topic of interest. Once the unidirectional relationship is formed, the user sees the full content of the enterprise.

Google+

Challenged by Facebook's domination of the social media scene, Google developed Google+ as a direct

competitor. Google focused on developing a solution that was comprehensive in its capabilities and integrated with other products in its portfolio.

Google+ users can create sub-groups within their larger community of friends, utilizing the concept of circles for organization and privacy. In addition, Google+ is integrated with Google Hangouts for desktop video conferencing. The circles feature gives users the opportunity to define the nature of their relationships with others and determine what content is shared with participants in their circles. The concept of circles increases the granularity of the security and privacy model but adds to the complexity for end users as they interact on this platform.

To advance from social media advertising to social media engagement, your business needs to develop your tactical approach to harvesting social media content.

Google+ has roughly 500 million users. Their main audience segment is males between the ages of 25 and 34. The most active users are business owners, bloggers, writers, technologists, and photographers.

Twitter

Twitter is a micro-blogging service individuals use to send “tweets” (messages) of no more than 140 characters. The size limitation forces users to be creative when posting messages. Twitter contributors often use links to more comprehensive sources of information. In addition, users employ abbreviations and modified versions of words to communicate using the fewest characters possible.

According to Twitter's site, at the end of 2014, the service had 271 million monthly active users and 500 million tweets were sent each day. Twitter's users are more active

than those of most social media platforms thanks to the ease of tweeting and the brevity of its content.

The Twitter “follow” relationship is unidirectional. Each user can choose whom they include in their feed by following them. If the followed user decides to follow back, the relationship becomes bi-directional.

With Twitter’s mechanism of “retweeting” content, the audience of a specific message can be extended to users outside the contributor’s immediate followers. This increases the viral potential of messages within the Twitter community.

Twitter is used by people of all ages and categories and it’s easy to tap into a portion of the community interested in a particular subject.

Pinterest

Pinterest was developed as a community of users who wanted to share content, mainly in the form of links, videos, and photos. It is often categorized along with other social bookmarking sites. It allows users to “pin” content onto their “boards” and share these boards with others. Much of the user base of Pinterest is women aged 35 to 50 who use the platform to share craft, décor, fashion, and cooking ideas with others who have similar interests.

The Pinterest “follow” relationship is unidirectional and selective. Users do not have to be fed all of another user’s content but can receive only the content associated with a specific topic or board.

Instagram

Instagram was born out of the idea that individuals are looking for an easy way to share photographs and want an easy method of applying filters to their photos to increase their level of interest to the community.

In Instagram, users follow others in a unidirectional relationship and can retrieve new photographic content as it is posted. Instagram has gained massive popularity, and although it is nearing 100 million users, the main influencers on this platform are celebrities and those in the 16–34 age demographic.

Harvesting Details

Once your business understands the communities of interest you wish to target with your message, you must understand what can be harvested and how the harvesting is done. Most large social media platforms have created a set of application programming interfaces (APIs) to allow third-party developers to both access and generate content. Although the APIs are similar, they have differing levels of access and control around the content.

**Social media is by definition big data
—its expansive volume of data comes
in different formats.**

Social media data is by definition big data—its expansive volume comes in different formats. In addition, the data is added to the public’s consciousness at variable velocity. For specific topics, new social media posts could be infrequent or could come at such a rate that social media platforms cannot or will not provide a complete stream of all new posts related to a topic through their APIs. Each social media platform exposes different levels and types of tools and APIs to allow information consumers to harvest this big data.

Firehose

The firehose is a stream of all posts available from a social media platform. Twitter offers this option to a handful of certified third parties. The certification process includes an extensive evaluation of the partner, including an evaluation of their functionality and technical capabilities, their customer references, and company financials. This ensures that these businesses can adequately serve as strategic partners.

Stream

A stream is a high-throughput, near-real-time access mechanism that can harvest data. These streams monitor a subset of the data based on parameters such as content hashtag, geographical location, or user. A stream is

similar to the firehose but has no guarantee that it includes every post, even within the defined subset. Depending on the volume of activity associated with the stream parameters, it might only provide a sampling of the data. Both Twitter and Instagram offer streams of content through their APIs.

A stream involves opening a communication channel with the platform and having content pushed as it arrives. It does not require a continual re-polling of the platform to request new content.

Search

The remaining social media providers can be searched for specific content through their APIs. The extent of the available content differs by provider and is bound by a provider's privacy and security model.

Twitter allows for search across all recent, publicly available content. Publicly available content includes all feeds that have not been defined as protected. Twitter is limited in the history it can search and weights results to focus heavily on recent posts. In addition, user profiles can also be searched.

Google+ allows search across all content that has been flagged specifically for public consumption. If content is flagged to be viewed by specific circles, it is not available as part of the public search. Google+ also allows for user profile search.

Facebook's core is the social graph, which includes the people and each of their connections. Facebook provides limited ability to search this graph. The major limitation on the Facebook API is the complex privacy model that Facebook has implemented to protect personal data and to give individuals control over who can access their content. In their current API, searching the content in posts has been deprecated. Facebook continues to support search capabilities to find users, events, pages, groups, and places.

Pinterest allows access to the site's top repins, most recent pins, related pins, most clicked-through pins, and pins resulting from specific search terms.

Instagram allows for searches of media by location and for user profiles. Like Twitter, it is limited in the history it can search.

Like streams, searches can include only a sample of the full content based on the total volume of the results of the search. If the volume associated with the search is immense, the likelihood of search results only including a sample of the data is much higher than if the search is highly targeted and has a relatively small result set.

Big data analytics can provide significant value by identifying context around independent elements of information, making them actionable.

Data Aggregation

With the diversity and complexity associated with different privacy models and API capabilities, third-party data aggregators have partnered with these social media platforms to provide unified streams of data for a price. Companies such as Gnip and Datasift allow companies to buy access to raw streams of social media data that have been aggregated from multiple social media sources. The companies also offer data enhancements to provide additional information that can be derived from the raw content through the use of analytics.

Content Generation

The real power of the social media APIs is their ability to interact with the social media platforms by posting content. The business model of social media platforms is based heavily on the ability of individuals to create content and share it with others. The more they can drive new, fresh, and relevant content on their platforms and increase overall site "stickiness," the more successful they are. For this purpose, this is an area of focus and interest for all social media platforms and their APIs.

As part of a social media strategy, the capability to post content enables the automation of social media advertising and engagement.

Strategy Step 2: Isolate Important Conversations

By utilizing a combination of near-real-time streams of information and focused searches, it is possible to amass a collection of content associated with any topic of interest. Harvesting information and storing it in its raw format amounts to creating a data lake. In this raw form, it has little value and will stagnate quickly.

In this situation, big data analytics can provide significant value by identifying context around these independent elements of information, making them actionable. Each morsel of information has additional context embedded in it that is not apparent at first glance. Within the content is information related to the segmentation of the user who produced the content and whether that social media user is a potential customer, the association of the content with products or services that the business is offering, and the sentiment of the content creator. Each of these information types is critical to determining whether further engagement with the content's creator is justified.

Market Segmentation

Market segmentation is of the utmost importance if the target objective of the business and its social media strategy is to identify and attract new customers. It allows a business to identify which users on the social media platform have similar interests to its existing customer base.

Each content creator has a personality, a set of preferences, and topics of interest. These are often manifested by a combination of their profile and the history of their posts.

One user profiling mechanism is to leverage concepts from information retrieval, where document similarity is computed by breaking down each document into a term vector and then calculating the algebraic resemblance to other term vectors or documents. Where each term vector has magnitude and direction in n-dimensional space,

a similarity score is based on the angle between these vectors.

This mechanism also can be applied to social media users and customers. A user's term vector is composed of the aggregation of their user profile content and the historical content of their posts. This term vector is created by decomposing the content into individual words and filtering out words that are highly common (called stop words) that act as noise masking the true meaning of the content. The individual terms users include to describe themselves in their profiles and the terms they use in their posts create a digital fingerprint of who they are.

Instead of engaging with every user who posts relevant content, a business can selectively engage with those users on the social media platform who are the most likely to become customers.

This alone is indicative of a user's preference and allows users to be segmented by commonalities but is magnified when a company has similar digital fingerprints for existing customers. By comparing the incoming user vector with a set of known good customer vectors, it is possible to identify new users who fit similar patterns of existing customers.

Instead of engaging with every user who posts relevant content, a business can selectively engage with those users on the social media platform who are the most likely to become customers.

If the objective is to enhance customer service and all user posts are considered to be existing customers, the mechanism associated with market segmentation is slightly modified. Instead of determining whether a

customer is closely related to an existing customer, the goal is to create a segmentation of known customers with known attributes and behaviors and then use the term vector comparison to identify the segment into which the unknown customer falls. This allows the engagement to be tailored to respond most effectively to a particular market segment.

The corpus used for sentiment analysis can be a generic set built on industry standards or it can be defined specifically by the business.

A classic example of this was when marketers looked at attributes of their customer base and identified a subset they called “soccer moms.” Realistically, not all individuals in this market segment had any interest in soccer, but there was homogeneity in attributes associated with their activities, the vehicles they drove, and their demographics (where they lived and their income levels). Identifying a social media user as a “soccer mom” allowed an enterprise to select how to manage the user (from a customer service perspective) and determine how to actively market to them.

Product or Service Fit

When harvesting content to attract new customers, providing personalized recommendations to a user in a timely fashion is an effective method of engagement. To accomplish this, your enterprise must identify the right service or product that responds to the momentary situation. A business that can match the right person with the right product or service at the right time enjoys an ideal scenario to create a positive first engagement. Such an engagement has been referred to as a digital business moment, or a transient opportunity, that can be exploited dynamically.

To accomplish this, associate key terms within the content with the relevant product or service. This can be

accomplished by creating a collection of trigger words for each product or service. When trigger words are included in social media content, the product or service can be identified and recommended to the content creator.

Take a maker of an effervescent tablet targeted at heart burn, upset stomach, and indigestion. When someone posts about discomfort caused by one of these target symptoms, a quick response from the business with a recommendation for their product provides timely and actionable information that offers the user a path to relief.

When addressing areas of customer service, understand what product or service relates to the customer’s post. Content is not always direct in its reference to products or services and only describe the nature or use of it.

To get from a description of the product or service to the actual product or service, a technique similar to market segmentation can be used. Term vectors associated with products can be compared to term vectors associated with the products and services to find the highest degree of similarity. If the business has a large product catalog, you can pre-filter products by comparing key terms in the content to key terms associated with the products.

Sentiment

With customer service, you can calculate sentiment for each post to understand the type of customer each is and what product or service a comment on social media is referring to, as well as the tone and attitude of the post. This is accomplished by developing a corpus of terms and determining their relative positive or negative sentiment value based on the overall tone of a known set of historical statements that include the same word. When the terms in a new post are scored using this pre-defined corpus, the sum of these scores provides an estimate of the sentiment of the new content.

The corpus used for sentiment analysis can be a generic set built on industry standards or it can be defined specifically by the business. If a business has product reviews with associated scores, it can take those scores and the terms associated with the product and develop its own sentiment corpus specifically for its product set and business model.

Strategy Step 3: Engagement

With a harvest of content that has been analyzed and enhanced to reveal previously unknown patterns, the final step is to use this information to instantiate an engagement with the social media end user. Several mechanisms are associated with engaging a user over social media: reply, mention, direct message, and targeted hashtag.

Reply

Most social media platforms have a mechanism for responding to the initial content. This can be in the form of a comment on the content or a reply to the post, allowing return messages to be sent to the initiator. If your objective is to provide additional information to the user who initiated the conversation, replying allows a message to be communicated back directly to the content originator. One advantage of this method is that the response is seen by the user as well as by their community of followers and friends.

In addition to active reply, there are more passive responses to a user and their content. These include following the user; re-tweeting, sharing, or pinning their post; setting as favorites; or “liking” the content. In conjunction with market segmentation, product identification, and sentiment analysis, these actions can communicate an intrinsic message to the user who created the post that their message was heard and the business agrees.

Direct Message

Another mechanism available on many platforms is a private direct message that is only seen by the target user. This provides the same opportunity as a reply but cannot be seen by others. This method provides a more intimate conversation, but its limited scope can have less impact because it's private. With some social media platforms, direct messages are only available if a bi-directional relationship is already established.

Mention

In addition to a reply, mentions can be used to attract a user's attention in a less intrusive manner. Including a user's name or username in marketing content is not a direct reply to a specific message but provides an

alternative mechanism acknowledging them as part of the message. This mechanism does not work in all platforms and can be construed as disconnected from the original conversation.

Targeted Hashtag

The basic organizational taxonomy in social media is the hashtag. A business can post a message with a targeted hashtag and insert that content into the general discussion about the topic. This is the least intrusive but is much less targeted at specific users with specific needs. Using hashtags focuses on communicating a message to the community as a whole at the same time that others are having a similar conversation.

Identifying the right social media users, the right message, and the right mechanism to deliver your message is doable, but winning hearts with a 140-character tweet or an Instagram picture is difficult.

Obstacles to Success

Social media is a great venue for achieving your marketing goals. Whether used for attracting new customers or caring for the existing customer base in a new way, social media has the potential for radically transforming the way businesses operate. With all of its advantages, it is not devoid of challenges.

Building Trust

Marketing is fundamentally about trust between your business and your current or potential customers. Identifying the right social media users, the right message, and the right mechanism to deliver your message is doable, but winning hearts with a 140-character tweet or an Instagram picture is difficult. In the end, analytic tools make it possible to home in on identifying and taking

advantage of a digital business moment, but remember that the goal is to build trust and confidence in your business, products, or services, and enhance users' perceptions of your business's ability to meet their demand.

Social Media Spam

As with other marketing mechanisms, generating unwanted engagements with customers may be viewed as spam. With the openness of APIs and the proliferation of spambots that send out large amounts of (unwanted) content, social media platforms are particularly sensitive to maintaining the quality of their streams of conversation.

Be sure to monitor user account behavior. Social media platforms are constantly looking for ways to identify and deter spambots from cluttering their communication stream. Platforms look for tell-tale signs of bot behavior by analyzing the frequency of posts, the redundancy of the content, and the size and nature of their users' friends and followers. Each platform can suspend or shut down an account that violates its acceptable-use policies.

Creepiness

Inevitably as messages are tailored to individual preferences and behaviors, some users will question how the business knows so much about them. Many will even feel such communication is an invasion of their privacy. The application of analytics to target the message to the right person at the right time can produce a degree of "creepiness." Such user unease becomes especially strong when more than one source is utilized to develop the digital fingerprint of the user. When a user profile on one social network is associated with a user profile on another social network, the picture of their personality and interests becomes clearer, but so does the potential intrusion into their personal lives.

To mitigate such user apprehension, let the customer opt out of having their social media traffic analyzed in an attempt to provide them a personal experience.

Resource Availability

As the role of big data analytics becomes fundamental in many facets of how businesses function, the demand for

individuals who can harvest and analyze big data intensifies. Resources with skills in harnessing APIs as data sources, consolidating information centrally, enhancing raw information with hidden attributes (such as sentiment and intent), and targeting the right message to the right user using analytics are often called data scientists.

These resources can be difficult to recruit and attract and are constantly being lured to other opportunities. One key to both attracting and retaining high-quality talent to support a business's social media strategy is to have a clearly defined purpose and objective and instill a passion for these into the people who support your enterprise's goals.

As social media becomes more pervasive, it becomes an important channel for communicating with customers. To be effective, a business needs to develop a strong social media strategy that is enhanced with big data analytics. The key is to target the right message, to the right audience, in the right digital business moment. When these attributes of the experience come together, social media advertising takes on the elevated role of social media engagement and makes experience more personal and intimate. ■

References

- Herschel, Gareth, Neil Chandler, and Jenny Sussin [2014]. *Hype Cycle for Customer Analytic Applications*. Gartner.
- Schaefer, Mark [2014]. *Social Media Engagement Is Not a Strategy*. Retrieved Oct. 15 from <http://www.slideshare.net/Rialta/social-media-engagement-is-not-a-strategy-36524665>