Deloitte.

Trusting big data: Perspective on data governance as a customer analytics investment

Deloitte Analytics

Many companies are investing significant amounts in customer analytics to drive their business and seek new ways to offer value to their customers. However, much of the potential value of that investment is at risk because data governance practices have not kept pace with the ways in which data is being used.

Defining the governance objective

A deluge of data now streams from nearly every aspect of your customers' daily lives and the interactions you and other business partners and data sources have with those customers. The opportunities to capture, understand, and gain insight into customer behavior and sentiment are immense. Many consumer-facing organizations are now investing in, accumulating, and analyzing their customer data. This significant increase in data use offers great opportunity, but also exposes organizations to remarkable risks.

There is no shortage of opportunities where data can be used, however, as businesses capture consumer activities, behaviors, and patterns, they are challenged with balancing two, potentially competing, objectives in the context of creating a trusted customer analytics program; how to simultaneously leverage and protect the data.

Behavioral analytics is a popular big data tenant helping businesses understand and predict likely customer demands and trends. It often involves multimillion dollar projects which gather a broad array of customer signals, and when successful, can be employed to increase brand awareness, improve product designs, promote loyalty, and enhance customer satisfaction. These insights can play a major role in defining product strategies and future business investments, and as such, are coming under greater scrutiny regarding their precision, inherent bias, and longer-term predictive confidence. The speed with which marketing decisions need to be made now demand that source data be collected, analyzed, and adapted in near real time.

Furthermore, as companies expand their co-marketing efforts or consider opening their coffers to business partners, they need higher levels of trust in their datasets and require confidence in their data governance framework. Companies that are able to convey assurance to this extended enterprise can gain a competitive advantage with a trusted customer analytics program.

The other concern, in addition to the reliability of customer insights, is how to respond to consumers' heightened awareness and concern about their privacy. Simply complying with the legal terms and conditions from a privacy policy no longer seems sufficient. Regardless of what is stated in that privacy policy, customers believe they have a "social contract" with the brands they trust. This social contract often contains expectations that go beyond the legal agreement and is the basis for how customers judge a company's behavior in the court of public opinion. The legal contract is of little defense when trying to win back customers that no longer trust your brand.

To create a trusted customer analytics program, data governance over customer analytics should seek to achieve two objectives:

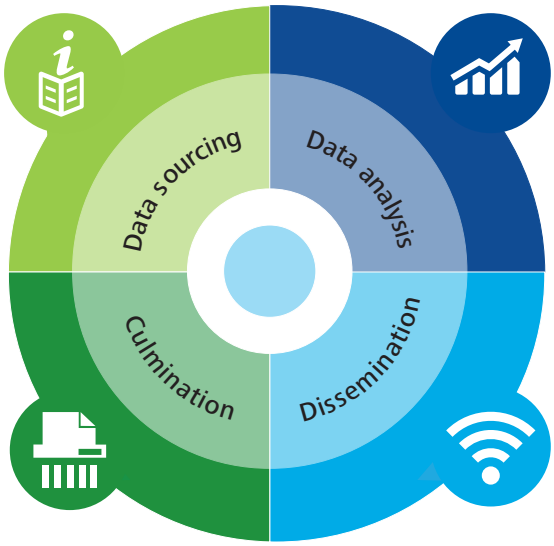
- reliability and precision of the source data and analysis performed on it.
- protection of the data to satisfy both customer and legal expectations.

While both objectives need to be achieved, there is a natural tension between both. More intense and widespread analysis of the data may lead to poor data-handling practices, placing customer privacy at risk. By the same token, too much control and rigidity can inhibit the ability to find key trends and insights. A strategic, proactive, and balanced strategy is needed.

The way forward

The quality and reliability of differently sourced datasets and the degree to which attribution can be made to individual customers varies significantly.

A helpful way to think about how to achieve the objectives is to break them down into logical phases of the data life cycle. Recognizing that your customer data needs to follow a structured data life cycle is the first realization for many companies in establishing a suitable control environment. We will explore the leading practices at each phase: data sourcing, analysis, dissemination, and culmination.



Data sourcing

Customer data signals vary widely in their source and reliability. Some data originates from internal enterprise systems, such as customer relationship management and billing systems, and have a higher degree of precision, structure, and security. Companies also generate data from marketing campaigns and in other cases acquire data from third-party sources. Data from internal systems tends to be structured and organized in accordance with an accepted data dictionary, whereas data from third parties (typically sourced from Web marketing and social media) tends to be unstructured with very few universally adopted standards defining how such data should be measured and reported. Other data that is purchased from third parties, such as the services provided by data aggregator brokers or social media listening platforms, are often sourced from disparate and sometimes unknown sources and have uncertain levels of reliability. So many customer analytics programs have a variety of data types from varying sources and are challenged with regard to how to combine and analyze these datasets. The quality and reliability of differently sourced datasets and the degree to which attribution can be made to individual customers varies significantly. More importantly, companies combining datasets from third parties should be asking the following questions:

- What methodology was followed in sourcing this data?
- Is the analysis and aggregation performed on the data certified and by what governing body or standards?
- Was the data ethically sourced (i.e., not obtained or sold under false or misleading pretenses and in compliance with relevant privacy regulations)?
- What resources, infrastructure, and tools are needed? big data toolsets today require specialized skills and the corporate infrastructure may need to be reviewed to manage the volume and disparate nature of the datasets.

Many aggregators and users of aggregated data are simply not asking these "trust" questions. The risks of not doing so are clear. The reliability of the analytical frameworks employed by a company based on external source data can be reduced to the reliability of the lowest-quality dataset, and the company runs a reputational and possibly a legal risk procuring unethically sourced data.

Companies should consider establishing a formal, enterprise-level data acquisition and ingestion program. This program should:

- Establish a data definition for all key data elements based, where possible, on industry-accepted standards.
- Perform due diligence on potential data suppliers to evaluate their data collection and protection practices.
- Evaluate the extent to which internal and external data sources meet the data definition standards.
- Develop a mechanism to tag data with respect to its relative reliability and sensitivity. This tagging is important to controlling data use in subsequent phases.
- Develop data analysis rules so that data is used in a manner consistent with its relative reliability and sensitivity.
- Regularly refresh the assessment of sources and datasets and adjust accordingly.
- Update infrastructure needs and data toolsets to reflect the ability to process and manage disparate data sources.

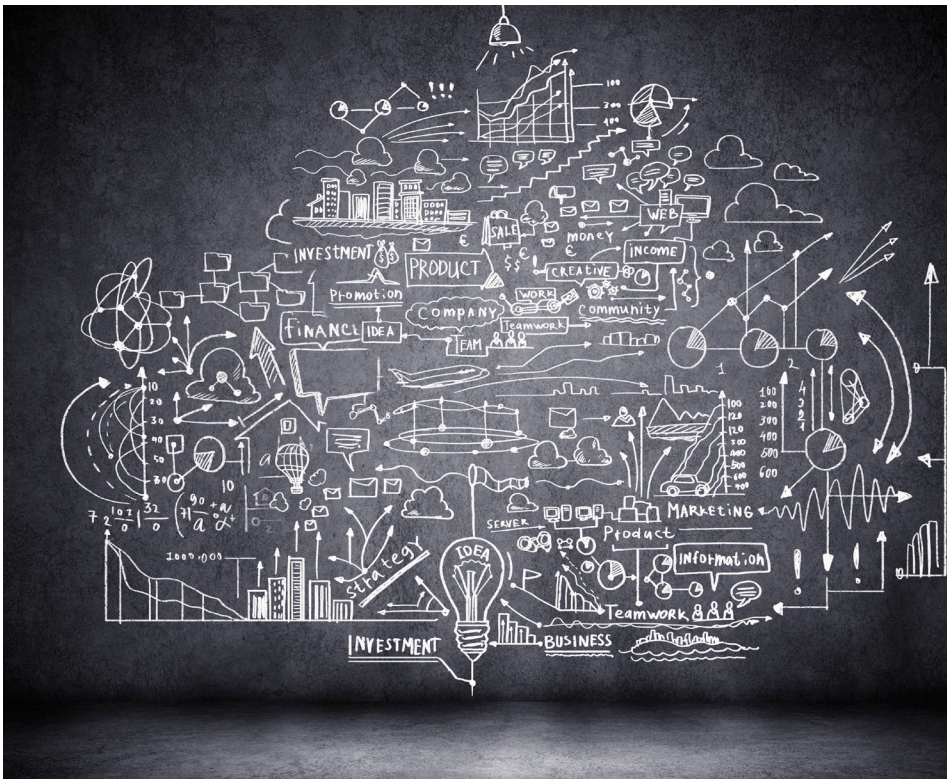
Such a sourcing program can take time and effort, but is critical to establishing a strong control foundation to the customer analytics program. The standards that are set for data sourcing will need to be reassessed periodically as customer profiling practices and data privacy regulations evolve.

Data analysis

There are a number of unique challenges when aggregating and analyzing customer behavioral data. Some of the more difficult challenges are in the variety of unstructured data types, including:

- Identification and removal of inherent bias.
- Attribution of marketing efforts to customer action.
- Estimation and forward protection of historical trends.
- Data precision and reconciliation.

As there is a degree of judgment and uncertainty in these types of analysis, it is important to identify these risks and define acceptable rulesets so there is consistency in practice and clarity on the risks inherent in the analysis. Unfortunately, it is still common in practice to see these risks overlooked or suppressed and an implicit assertion that the reporting coming out of customer analytics is both reliable and trustworthy.



Dissemination

Controlling the dissemination of reporting is important not only to protect your customer data and intellectual property, but also to ensure the reporting is used by those authorized and for those purposes for which it was created. Business groups will often request open access to customer analytics data stores to have the freedom to do their own analysis and to not be "shackled" by the formality and time it takes to produce formal reporting. The danger, however, is that it can be very difficult to reign in excessive access once it has been given.

Leading practices, tools, and infrastructure in the dissemination of customer analytics reporting involve controlling access both from the perspective of who should see the reporting as well as controlling which reports can be accessed. For example, some reports that contain time-sensitive data should be withdrawn when they are no longer reliable. Another leading practice is clearly describing in reports the assumptions used, risks thereof, and purpose for which the reporting has been prepared.

Looking forward, companies that are in the business of selling customer data or insights, or, those that need higher levels of internal confidence, should consider establishing a framework to convey assurance to users. Such a framework should cover the integrity of the reporting and the protection afforded to the customer data used in the reporting. The Service Organization Control (SOC) reporting framework issued by the American Institute of Certified Public Accountants (AICPA) is one example of such a framework that could be used to convey such assurance.

Culmination

At some point, data will outlive its usefulness and a leading practice is to be intentional and set policies for the time period various types of customer data will be used and ultimately retained. There are a number of reasons why companies should be intentional about data destruction. The utility of customer data tends to erode over time and the longer it is retained the greater the risk it presents for drawing conclusions. In addition, the risk associated with a breach of customer data also increases if more data is retained than is needed. The larger and older the customer dataset, the more difficult it is to track and trace to the customers it relates to, and manage in a breach situation. Finally, the legal risk of complying with electronic discovery requests becomes much more difficult, if not impossible, to comply with unless these types of requests are anticipated in advance. Some data will need to be retained for longer periods for various reasons but too often we see customer analytics practices that are either unintentional about data destruction or have a mind-set that all data should be retained forever.

Leading organizations have implemented a robust data governance framework that assigns accountability for the tracking and inventory of customer datasets. Included in these frameworks are defined data policies that require data be destroyed after a certain period of time. For sensitive data, it is important to be able to confirm the destruction of such data. The certification of destruction should consider all repositories such as off-site backups and should use technologies that eliminate the ability to restore such data.

Conclusion

With the investment in customer big data programs growing, the potential value add for companies that get this right is significant. However, success demands a robust control environment. Companies that underinvest or delay their investment in their control environments over customer big data, threaten that investment. Aside from the fact that a strong control environment has significant benefits, rising expectations of confidentiality and security in contracts, law, and regulation, as well as the potential from brand damage when data is leaked or misused, are quickly making this a business imperative.

The companies that will ultimately succeed with customer big data are those who have a strategic and proactive framework that can get consistent and more reliable insights while carefully protecting that data.

For more information about data governance, analytics, and the risks to your organization, please contact:

Contacts

Tim Davis
Deloitte Advisory Principal
Deloitte & Touche LLP
+1 206 716 7593
timdavis@deloitte.com

Sunita Suryanarayan
Deloitte Advisory Principal
Deloitte & Touche LLP
+1 213 553 1339
ssuryanarayan@deloitte.com

Mohammad Nayeri
Deloitte Advisory Senior Manager
Deloitte & Touche LLP
+1 213 688 5532
mnayeri@deloitte.com

Glenn Wilson
Deloitte Advisory Manager
Deloitte & Touche LLP
+1 213 688 6976
glennwilson@deloitte.com

This publication contains general information only and Deloitte is not, by means of this publication, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your business. Before making any decision or taking any action that may affect your business, you should consult a qualified professional advisor.

Deloitte shall not be responsible for any loss sustained by any person who relies on this publication.

As used in this document, "Deloitte Advisory" means Deloitte & Touche LLP, which provides audit and enterprise risk services; Deloitte Financial Advisory Services LLP, which provides forensic, dispute, and other consulting services; and its affiliate, Deloitte Transactions and Business Analytics LLP, which provides a wide range of advisory and analytics services. Deloitte Transactions and Business Analytics LLP is not a certified public accounting firm. These entities are separate subsidiaries of Deloitte LLP. Please see www.deloitte.com/us/about for a detailed description of the legal structure of Deloitte LLP and its subsidiaries. Certain services may not be available to attest clients under the rules and regulations of public accounting.

Copyright © 2015 Deloitte Development LLC. All rights reserved.
Member of Deloitte Touche Tohmatsu Limited