

Analytics Data Lake

Using Big Data to Successfully Drive Customer Engagement

Bridging Data Siloes

One major driver of personalized customer engagement is analytics across all types of customer data – including demographic, transactional, and behavioral data. Understanding customer behavior requires three steps: Collecting and updating all relevant data; building an accurate customer profile or Golden Record; and analyzing the data – including detailed current behaviors – to understand the customer’s journey and determine next best actions. Using a data lake as the basis of behavior analytics is an intriguing idea, but without the right approach, a data lake can become a treacherous swamp.

Organizations have struggled for years to access the customer data needed to provide personalized and contextual experiences across all touchpoints. A data lake is a cost-effective way to capture the detailed data associated with consumer behavior, e.g. streaming device data or website behavioral data. Many solutions have focused on a self-service approach to access the data in the data lake.

Both storage and access are necessary, but far from sufficient. According to a recent Forrester report, “only 25 percent of business and technology decision-makers report seeing increased revenue from their implementation of big data solutions. This means that a clear majority of companies are not effectively harnessing the insights in this data to win, serve, and retain their customers.”

So, what’s inhibiting successful use of the data lake in customer engagement? Two technology bottlenecks stand between the data lake and using analytics for customer engagement: 1) creating the Golden Record – the accurate and complete view of the customer – and 2) overcoming latency within the process (at the data, analytical, and execution levels) to engage with the customer in real time on the correct channel or touchpoint.

The RedPoint Customer Data Platform™ (CDP) eliminates data lake bottlenecks and delivers cleansed and accurate customer data to analytics and execution systems while leveraging existing Big Data technology. The CDP can power legacy systems in place, and ensure data is prepped and available for model building and deployment in RedPoint AI Studio™ and RedPoint Customer Engagement Hub for comprehensive enterprise customer experience optimization.

Big Data – Promises and Problems

A data lake is a manifestation of the Big Data movement, using Apache Hadoop to provide clustering and file storage, and a set of (mostly Open Source) database formats and processing tools to bring the data lake to life.

The promise of Big Data is to handle all data, existing and new, without the need for IT to reformat everything and place it into standardized SQL databases. Data Lakes address part of this promise by offering space for files and information in multiple different formats, along with toolsets to allow access and analysis of the information in the data lake. Additionally, many big data technologies (such as NoSQL databases), offer “schema-on-read,” where the schema is created when the data is read vs. when it is written. This requires less IT involvement in database design but may make it extremely difficult for business users if data quality is not built into the process.

Big Data is characterized by “Three Vs”: Volume, Variety, and Velocity. More recently, additional “Vs” – Veracity and Value – have been added as dimensions to consider. A customer analytics solution – based on a data lake or not – must be able to cope with the volume of existing data and the velocity of new data. It must also be able to read and correlate information across a broad variety of formats.

A data lake offers the framework for handling the core “three Vs” on the intake side but needs additional “intelligence” to evaluate data for accuracy, consistency, and timeliness (“Veracity” or truthfulness of the data) and to cleanse, correlate, and analyze the data, activating it by presenting data and analytic insights to the systems and people engaging with customers. Without timely activation based on accurate data and relevant insights, data quickly loses value, becoming irrelevant to customer engagement.

RedPoint CDP performs data ingestion and integration across all sources and types of data in the data lake – batch or streaming, structured or unstructured – to provide an always-on, always-updating customer profile, i.e., a Golden Record. RedPoint CDP is certified on all cloud infrastructures, Hadoop distributions, and leading databases.

RedPoint CDP and Data Lake Technology

The RedPoint Customer Data Platform works with a variety of data storage technologies, though Hadoop is commonly used as it is a cost-effective technology. There are three methods used by RedPoint in working with Hadoop specifically:

Reading from and writing to the HDFS file system: This allows the CDP to ingest HDFS-based data into RedPoint CDP's sophisticated data transformation, cleansing, and matching algorithms. RedPoint lets users easily combine data stored inside and outside of HDFS, and enables high-performance extraction, transformation, loading, and format conversion between HDFS and external sources.

Reading from and writing to Hadoop and Cloud Databases: RedPoint CDP provides direct connections for a broad set of Hadoop and Cloud database types, including:

- Google Storage
- Avro
- Parquet
- HBase
- Hive
- Google BigQuery
- AWS S3
- AWS Redshift
- AWS EMR
- Azure HDInsight

Processing inside Hadoop as YARN tasks: RedPoint uses Hadoop's YARN management system to create and run a series of parallel tasks to carry out its work. RedPoint projects are distributed to worker nodes based on tasks to be performed, data requirements, and HDFS data locality.

RedPoint's architecture massively scales complex application-driven processes, such as data cleansing and hygiene, mapping and GIS functions, householding and advanced matching. As RedPoint software works natively in Hadoop, there is no need to move the data outside of Hadoop to perform data quality functions. All access, cleansing, and enrichment functions can be done within the Hadoop cluster, further resulting in time compressions up to 99% in moving from data-to-insight-to-action.

Business Value Delivered

The RedPoint CDP acts as an intelligent agent for a data lake, improving the speed, accuracy, and relevance of analytics and engagement initiatives by producing complete and up-to-date customer profiles, the "Golden Records" needed to help data analysts discover, enrich, and deploy relevant models and views of customer information.

Combined with additional RedPoint software solutions for advanced analytics and omnichannel orchestration, enterprises can achieve superior business benefits. RedPoint's CDP also ensures data scientists and analytics professionals will spend less time on data prep and more time on strategic model development and deployment. The Golden Record generated by the RedPoint CDP can be leveraged for up to date insights that inform ongoing omnichannel customer engagement strategies. Specific benefits businesses can expect to see include:

- **Reduced operational expense and increased productivity:** Marketers can more easily access data and models to optimize journeys. Data scientists spend less time managing data and more time doing analysis.
- **Increased personalization and relevancy at scale:** Combine personalization with real-time capabilities to increase revenue by delivering messages and offers in the cadence and context of the customer journey. More relevant messages lead to less customer fatigue, friction, and lower interaction costs.
- **Increased customer lifetime value and revenue:** The robust path-to-purchase model scores each customer on his/her purchase journey, which lowers the cost of both acquiring new customers and retaining existing loyal ones.

About RedPoint Global Inc.

RedPoint Global's software solutions empower brands to transform how customer experience is delivered. RedPoint Global's solutions provide a single point of control to connect all customer data, determine next best actions in real time, and orchestrate interactions across all enterprise touchpoints. Leading companies of all sizes trust RedPoint Global to deliver highly personalized and contextually relevant experiences that optimize customer engagement. For more information, visit www.redpointglobal.com/aistudio or email contact.us@redpointglobal.com.



US Headquarters: Wellesley, MA | Tel: +1 781 725 0250 EMEA Headquarters: London, UK | Tel: +44 (0)20 3948 8170
www.redpointglobal.com