

A Guide to Selecting the Right Customer Data Platform (CDP)

Digital transformation is profoundly impacting practically every business today. It is commonly understood that customer engagement is one of the key foundational capabilities underpinning this transformation.

It is also understood that the ability of an enterprise to effectively engage its customers is dependent on its ability to create and operationalize the most complete and accurate view of a customer and make it available for use at every customer touchpoint. Customer data platforms (CDPs) are a type of data platform required to operationalize all data about customers at the speed, accuracy and depth required to effectively drive transformative levels of engagement.

The objective of this CDP Guide is threefold and provides:

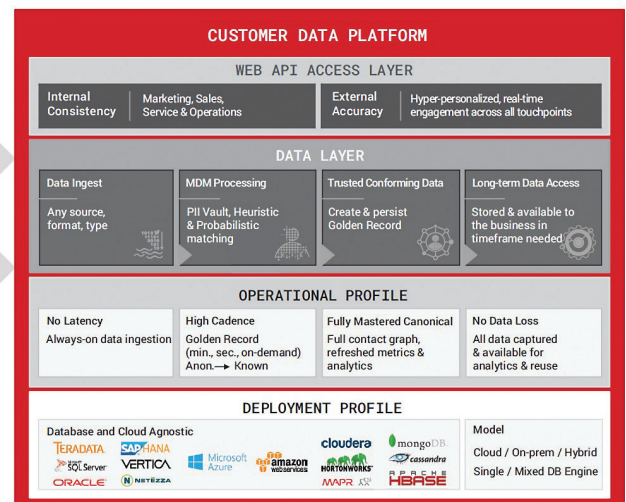
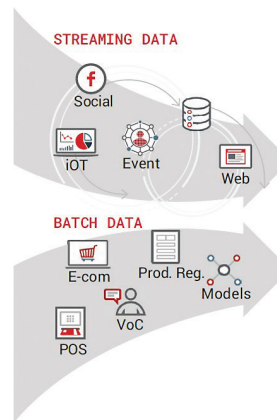
1. Redpoint Global's definition and point of view of a CDP.
2. Various industry analyst CDP definitions and related criteria to help organizations understand how to select the CDP that best enables digital transformation.
3. Definitions and shortcomings of alternative technologies and techniques that are not CDPs.

Key performance characteristics of the Redpoint Customer Data Platform include agility, precision, scale, speed and accessibility needed to be the data engine driving enterprise digital transformation:

- **Agility** – Support for all data sources and unparalleled flexibility: The Redpoint Customer Data Platform is architected to work with all sizes and types of data whether structured, semi-structured, or unstructured data. It also provides out-of-the-box connectors to any environment spanning traditional databases, applications and advanced Hadoop/data lake and other No-SQL environments. The platform integrates first-, second-, and third-party data to create a complete and always fresh view of a customer that can be leveraged in a real-time cadence for analytics and decisioning. There are no limits on the types of data or data structures, making it easy to add new sources at any time.

Redpoint's Definition

The Redpoint Customer Data Platform™ (CDP) is a new type of operational data environment that ingests an enterprise's data from all sources – whether batch or streaming, internal or external, structured or unstructured, transactional or demographic and personal or general – that provides an always on, always updating golden record and makes it continually available at low latency to all touchpoints and users across the enterprise. The Redpoint CDP is in production around the world delivering quantified value to a wide spectrum of enterprises.



- **Precision** – Advanced data processing and matching algorithms: Redpoint incorporates sophisticated data transformation features including advanced parsing, complex processing rules and computation, address and other data standardization, geocoding and spatial analysis. It performs deterministic and probabilistic identity resolution, persists unique superkeys and sub-keys and provides customizable levels of groupings including individual, household, segment and business.
- **Scale** – Production scale with drag-and-drop ease of use: Users can set up automated workflows with error handling alerts, checkpoints and restart mechanisms, version control and job monitoring to go from design to production with ease. Graphical representations make it easy to create, review and modify workflows without writing code. The Redpoint CDP integrates seamlessly with legacy and new enterprise systems including: operational systems, data lakes, business intelligence and reporting systems, and synchronous and asynchronous engagement systems.
- **Speed** – Real-time data ingestion and integration: Redpoint processes data at a speed and scale that's superior to any other provider in the market, whether batch or streaming. Data loaded into the CDP is immediately available for consumption, whether that is for analysis, decisioning, or to drive operational systems. Benchmarks show that Redpoint is 5x-20x faster than other solutions evaluated by industry analysts. Redpoint uniquely does this through native data access that reduces latency across all data processing topologies.
- **Accessibility** – All business applications in the enterprise can access the detailed and current canonical data through a streamlined services layer, at speeds that support real-time operational use. Business users can curate data through master data management (MDM) processes and user interfaces that place the data closer to the point of need.

Forrester's Definition

Forrester published a definition in November 2016 of a type of CDP they refer to as big data fabric, which "offers enterprise architecture (EA) pros a platform that helps them discover, prepare, curate, orchestrate and integrate data across sources by leveraging big data technologies in an automated manner."

In the *Forrester Wave – Big Data Fabric, Q4 2016*, Forrester outlined four high-level criteria to access and leverage big data with a modern architecture:

- **Delivers new actionable insights with minimal effort** – Big data fabric offers the ability to aggregate, transform, cleanse and integrate data from multiple big data sources, to deliver insights with zero to minimal coding.
- **Secures big data end-to-end** – Big data fabric enables centralized data access and control.
- **Enables real-time integrated data across the business** – Big data fabric enables data and metadata sharing between peers, employees, partners and customers. It allows any application, process, dashboard, tool or user to access any integrated data, regardless of where the data is physically or logically located and regardless of the data format.
- **Delivers a self-service data platform for business users** – Big data fabric emphasizes self-service data preparation, curation, orchestration and integration services that nontechnical personnel can leverage for improved decision making.

Redpoint's CDP meets all of Forrester's criteria in several innovative ways:

Forrester Criteria	Redpoint's CDP Provides...
Delivers new actionable insights with minimal effort	...drag-and-drop tools that enable businesses to integrate, transform and match data with no programming required, and has proven to save over 80 percent of the manual effort in preparing data for advanced analytics while delivering higher quality insights.
Secures big data end-to-end	...a single customer view or golden record, underpinned by a persistent superkey and its sub-keys, is accessible across the enterprise, while providing business users the tools for master data management (MDM) so that the business is in charge of the customer data, e.g., defining the matching rules and tolerance levels to automatically de-duplicate data, and providing data stewardship and governance needed to truly drive enterprise-wide customer engagement.
Enables real-time integrated data across the business	...the golden record is updated in real-time, incorporates real-time identity resolution and is accessible across the enterprise to fuel real-time analytics, next-best-actions and customer engagement systems in real time.
Delivers self-service data platform to business users	...data preparation, curation, orchestration and integration services through drag-and-drop tools so that businesses are closer to the data used to drive analytic models and highly automated and personalized customer engagement; MDM capabilities also put data closer to the edge where it is used, enabling business functions to master customer data with easy to use but powerful MDM functionality.

Gartner's Definition

Gartner published a report focused on CDPs in November 2016 that defines the CDP as "an integrated customer database managed by marketers that unifies a company's customer data from online and offline channels to enable modeling and drive customer experience." Gartner also outlined CDP selection criteria in this report, which was titled *Innovation Insight for Understanding Customer Data Platforms*:

- **Identity** – How is customer data linked, persisted and refreshed, can it do deterministic cross-device matching and probabilistic matching.
- **Data connectors** – Assess which existing tools will be supported out of the box by each CDP. Consider how data is collected, whether via APIs, server to server, uploads, or tags.
- **Channels supported** – Some CDPs integrate data from channels better than others.
- **Data processing** – CDPs should feature batch and streaming data processing and offer greater scalability and flexibility than a traditional customer database.

61% of CMOs believe they are under-leveraging big data.

CMO Council

- **Segmentation** – Consider how the CDP handles segment discovery. Is it rule-based, or does it allow for statistical techniques, such as clustering?
- **Predictive modeling** – Most CDPs feature propensity models or lead scoring or allow analysts to incorporate their own models. Increasingly, CDPs are incorporating next-best-offer recommendations and other decision support features.
- **Orchestration** – Marketers will generally still need execution systems for the final mile, but CDPs will support cross-channel orchestration and personalization by managing rules, instructions and messages.

Gartner's CDP criteria is more expansive than other analysts, and Redpoint's technology effectively meets all of these criteria as well:

Gartner's CDP criteria is more expansive than other analysts, and Redpoint's technology effectively meets all of these criteria as well.

Forrester Criteria	Redpoint's CDP Provides...
Identity	...the most advanced algorithms in the market for probabilistic matching, as well as persistent key management, the ability to manage identities from unknown to known states through progressive profiles, and the ability to resolve identities at multiple levels (person, household, digital address, business, device); Redpoint has been ranked #1 for operational and transactional data quality and #1 for data integration in Gartner's Critical Capabilities for Data Quality Tools report.
Data connectors	...hundreds of out-of-the-box connectors to ingest data from any source, whether unstructured big data from a Hadoop cluster, or traditional structured data from a relational database; a software development kit (SDK) also provides ease of implementation for any customer connectors needed.
Channels supported	...an open garden architectural approach, so that data can be integrated from any engagement systems of record, whether proprietary or widely used third-party channels, social media networks, devices or any other touchpoint.
Data processing	...processing of batch and streaming data, at a scale 5x-20x faster than other solutions evaluated by industry analysts, providing the ability to operate at the speed of the customer; Redpoint has compressed the time from data-to-action by up to 99 percent for high volume clients.
Segmentation	...rules-based, clustering models and machine learning algorithms to segment the customer base, enabling enterprises to scale to a true segment-of-one while maintaining a clear point of operational control.
Predictive modeling	...predictive models and machine learning to optimize customer decisions, with the option to ingest predictive models from an enterprise's library, third-party libraries or Redpoint's library of advanced models.
Orchestration	...the underpinning of the Redpoint Customer Engagement Hub™, which provides in-line analytics and intelligent orchestration to provide a single point of control over customer next-best actions and decisions to personalize engagement across all systems of engagement.

CDP Institute's Definition

The CDP Institute (www.CDPInstitute.org) definition outlines that "a Customer Data Platform is a marketer-managed system that creates a persistent, unified customer database that is accessible to other systems." Redpoint's technology provides robust capabilities that meet the CDP Institute's criteria in advanced ways.

A CDP provides an always on, always updating golden record and makes it continually available at low latency to all touchpoints and users across the enterprise.

This definition has three elements:

CDP Institute Criteria	Redpoint's CDP Provides...
Marketer-managed system	...a customer view and MDM capabilities that can be managed by marketers, or managed as an enterprise-wide capability used across marketing, sales, service and all enterprise touchpoints.
Creates a persistent, unified customer database	...a single customer view or golden record, underpinned by a persistent superkey and its sub-keys, is accessible across the enterprise.
Accessible by other systems	...a services layer that provides real-time accessibility to any engagement system, analytic platform, or other application across the enterprise.

Ways That Alternative Customer Data Technologies Fall Short

There are other customer data technologies that sometimes get confused with CDPs, when they are at best complementary technologies. Simply put, organizations cannot build their CDPs on these technologies. Other customer data technologies such as data warehouses, marketing clouds, DMPs, data lakes and event hubs lack the keying, precision, matching, data connectivity, immediacy and cadence associated with a CDP that consumes and connects all data, all the time, and operates at the speed of the always-on customer. All other options have a distinct purpose designed to support a different business function, none of which by themselves can address the customer engagement demands in the digital world. Given the shortfalls of other technology options, it is no wonder that 61 percent of CMOs believe they are under-leveraging big data. In addition, only seven percent of marketers are able to deliver real-time, data-driven engagements across both physical and digital touchpoints (CMO Council).

Several approaches to building the unified customer database have been offered in the past, each with substantial drawbacks. All of these alternative customer data technologies fall well short of providing instant recognition of a person in a broader digital context, along with the capability to provide real-time access to a golden record that drives next-best actions in engagement systems across the enterprise. In most cases, CDPs play an important role in providing data to, and/or ingesting data from these technologies.

Data Technology	Why this technology is not a CDP...	How this technology can work with a CDP...
<p>Data Marts and Data Warehouses</p>	<p>Data warehouses summarize data for reporting purposes and lack the detailed data required to effectively personalize interactions. They are structured to analyze large data sets, not to utilize information about a single customer during a real-time interaction. New data is typically loaded at intervals, not in real time, creating significant delays in data cycle times due to reformatting, cleansing, reorganizing and indexing.</p>	<p>CDPs provide the data ingestion, cleansing, matching and identity resolution capabilities to connect data. This connected data can be loaded into a data mart or data warehouse to support reporting applications. CDPs maintain the detailed data that is lost in data warehouse summaries, and CDPs provide real-time unified customer views that are accessible by engagement systems enterprise-wide.</p>
<p>Marketing Clouds and Campaign Management Systems</p>	<p>Conventional campaign management systems were designed to connect to a separately built data warehouse and had few if any built-in features to maintain their own database. Databases are typically an afterthought for these solutions since most of them originated as list based systems for direct mail, blasting emails, or for other broad communications. Marketing clouds include a customer database, but most of these have a fixed data structure with limited size and flexibility. They lack robust data cleansing, probabilistic matching, linking of first-, second-, and third-party data, identity resolution, persistent key management and data curation and stewardship capabilities. Furthermore, they actually propagate data fragmentation and their "walled garden" approach limits the control an organization has over its own data.</p>	<p>A modern marketing architecture includes a CDP as the foundational data layer to fuel analytic models and engagement flows to deliver highly personalized and contextually aware interactions across all enterprise touchpoints. The CDP also provides a level of flexibility in data design and connectivity, control over data and a set of open APIs to connect to an organization's broad marketing stack whether on-premises, in the cloud, or a hybrid deployment.</p>
<p>Data Management Platforms (DMPs)</p>	<p>DMPs were originally designed for programmatic advertising and to store lists of browser cookies for advertising audience selections and real-time bid management. They are typically limited to a simple data structure that sharply restricts the amount of detail that can be stored about a customer and the length of time the data persists. DMPs lack identity resolution capabilities across first-, second-, and third-party data and cannot store any personally identifiable information. DMPs don't process data quickly and typically take between 24-to-48 hours to onboard data, which creates data latency problems for marketers when trying to engage customers in highly personalized and real-time terms.</p>	<p>DMPs are simply another interaction channel that can be integrated with a CDP. CDPs ingest data of any type from multiple channels, integrate it all together and allow access to orchestrate customer engagement and serve and enhance wider business functions beyond marketing. CDPs drive business users and enterprise applications that require data at any level of detail – across anonymous and known interactions – and no matter if it is online or offline engagement or "n" party data.</p>
<p>Event Hubs</p>	<p>Event hubs provide an on-ramp for streaming data. Given the scale of streaming data that they are tuned for, they completely lack the intelligence to effectively integrate and connect data. These event hubs capture the streaming data in a big data storage environment to be processed later.</p>	<p>CDPs are able to access the data from event hubs to connect it and drive insights from it. As these event hubs are built to capture the streaming data in a big data storage environment to be processed later, CDPs can then identify start and stop points in the streaming data and connect it to a broader customer golden record.</p>
<p>Data Lakes</p>	<p>Data lakes provide a lower cost option to store massive amounts of any kind of data, and the infrastructure to process data, but lack the application smarts to connect data and make it accessible in real-time. Data lakes are typically needed for big data, but insufficient in supporting business initiatives to optimize customer engagement.</p>	<p>CDPs use data lakes for the role they were built for – as infrastructure to process and store data – and CDPs go beyond this to provide intelligent linking of customer data or create a unified customer view necessary to guide highly personalized engagement.</p>

Glossary of Terms

Canonical Data – Data that is trusted and conforms to the accepted rules and procedures of an organization. Canonical data provides a level of consistency and accuracy such that it may be used across an enterprise.

Contact Graph – All of the proxy identities that may represent a person across all the touchpoints between a brand and a customer, e.g., email addresses, cookies, device IDs, social handles, physical addresses, phone numbers, user names, mobile app names.

Customer Data Platform – A customer data platform (CDP) is a new type of operational data environment that ingests the enterprise's data from all sources – whether batch or streaming, internal or external, structured or unstructured, transactional or demographic, personal or general – to provide an always on, always updating golden record and make it continually available at low latency to all touchpoints and users across the enterprise.

Digital Transformation – A fundamental shift in how enterprises provide digital engagement, products and services, along with a shift in how consumers connect to the people, brands and things that matter the most to them. From mobile apps to wearable fitness devices, WiFi accessible programmable home thermostats and smart refrigerators, consumers expect aspects of their daily lives to connect digitally to products and services in ways that are highly relevant while being seamless and without friction. Customer engagement in this digital era demands highly relevant and personalized offers, messages and content across all enterprise touchpoints.

First-, Second-, and Third-Party Data – First-party data is the data captured by an organization from a growing range of transactional and engagement systems that a brand owns such as CRM, eCommerce, customer service/call center systems, websites, loyalty applications, POS systems, etc. First-party data might include personally identifiable information (PII) which can be used to identify a specific person. Whether data about customers, members, patients, policy holders, accounts, or products, first-party data serves as the cornerstone of the golden record. Second-party data is data that is collected by business partners which may be franchisees, dealer networks, or other co-marketing partners and used in combination with first-party data. Third-party data is widely used by organizations to build richer, more detailed customer profiles, with the data provided by independent sources such as data aggregators and credit bureaus that may be appended to the golden record.

Golden Record – A record that combines meaningful, reliable data from multiple systems into a single, best record. It is the "ideal" of what a record can and should be and as such, it stands as the "truth." It is a combination of the most complete contact graph coupled with metrics, behavioral information, purchase history, customer value, model scores, and preferences that represent a person. The golden record is more accurate and more complete than the data from any single source. A persistent key is maintained that enables a

progressive profile, dynamically adding data to the golden record as customers move from an anonymous to a known state, and tracked across all stages of the customer journey.

Identity Resolution – An operational process to identify an entity – for example, a person, household, or device – through an automated process that may use a combination of deterministic and probabilistic matching. Identity resolution is the cornerstone operational process to drive data-quality initiatives, progressive profiles, and golden records. This operational process may include standardizing, normalizing, validating, and enhancing data as part of an automated process.

Master Data Management (MDM) – Master data management (MDM) is the technology and process to de-duplicate data, and resolve differences in data to capture the single version of truth. Data may be de-duplicated in an automated manner, and/or through data curation and governance tools placed in the hands of the business user that is closest to the use of the data.

Matching/Deterministic Matching – Linking records based on an exact or nearly identical match of identity attributes, e.g., name, address, customer number. These typically work in isolated and siloed data systems that have highly constrained data entry. This is also referred to as hard-key matching.

Matching/Probabilistic Matching – Using a wide range of identifiers, weights, and probabilities to correctly link customer records or other entities such as devices. These links can be enhanced with heuristic matching rules that learn and improve from experience over time. With advanced matching, organizations are able to prevent duplication, make connections, and identify matches to bring together various identity proxies (names, addresses, households, digital identifiers) to unify communications with their customers.

Operationalizing Data – Preparing data and making it available anywhere in the enterprise, at the cadence needed to drive results. To operationalize data is to deliver the data to where the value is created so that the data can be monetized, including delivering data to in-line analytics and to drive real-time customer engagement decisions.

PII – Personally identifiable information (PII) is information that can be used to identify, contact, or locate a specific person, and is bound by information security and privacy rules, regulations, and laws.

Sub-Key – A proxy identity, such as email address, social handle, name, address that may be linked to a superkey; though they may change, the sub-key and its links persist over time.

Superkey – The linkage between all of the potential proxy identities and behaviors of a person. A superkey maintains its stability over time, as sub-keys change. For example, if social handles or email addresses change, the superkey persists for that person and is updated with other data as it becomes available, maintaining a true and current single view of a customer.

A Roadmap for Moving Forward

Digital transformation is fast becoming a significant part of the business landscape today, but the disruption it will cause over the next three to five years is difficult to fully imagine. One thing for sure, however, is that none of it will be possible without a new level of fluidity and accessibility of clean, continually refreshed, connected, low-latency customer data, available on demand across the enterprise. A CDP is at the center of digital transformation and customer engagement strategies, effectively connecting data to compete in today's market. The deep customer understanding possible with a CDP enables brands to deliver consistent, relevant, and highly personalized customer engagement in today's digital economy.

Given the strategic importance of CDPs, it is critical to be clear on the capabilities needed to support a given enterprise's strategy, as vendors and technologies claiming to be CDPs will provide different ranges of features and performance. There is an emerging set of evaluation criteria that analyst firms such as Forrester and Gartner and the CDP Institute recommend when considering CDP capabilities, and they serve as solid guidelines for selecting the right CDP solution. Redpoint has sharpened these definitions based on its experience delivering CDP solutions at scale in the market today. Redpoint has delivered CDPs for innovative enterprises across practically all consumer-driven industry segments, including Financial Services, Retail, CPG, eCommerce, Healthcare, Insurance, Media and Entertainment, Travel and Hospitality to name a few. While CDPs may be an emerging category, the Redpoint Customer Data Platform has been proven at the scale, speed, precision, agility, and availability required by the most demanding digital transformation initiatives.

About Redpoint Global

With Redpoint's software platform, innovative companies are transforming their customer experiences across the enterprise and driving higher revenue. Redpoint's solutions provide a remarkably unified, single point of control where all customer data is connected and every customer touchpoint intelligently orchestrated. Delivering more engaging customer experiences, highly personalized moments, relevant next-best actions, and tangible ROI—this is how leading marketers lead markets. To learn more, visit redpointglobal.com.



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

© 2020 Redpoint Global Inc. All rights reserved. Redpoint, the Redpoint logo, and all Redpoint product names are trademarks of Redpoint Global Inc. All other trademarks are the property of their respective owners.