The Case for a Real Time Customer Data Platform

Raab associates, inc.
730 Yale Avenue
Swarthmore, PA 19081
www.raabassociatesinc.com
info@raabassociatesinc.com

The Case for a Real Time Customer Data Platform

Introduction

Customer Data Platforms have emerged in recent years as a way to help marketers build a complete customer view. They provide a faster, cheaper alternative to a conventional data warehouse while avoiding the lock-in of single vendor suites and marketing clouds. The latest twist in the CDP story is versions that support real-time interactions in addition to traditional batch approaches. This paper will explain what CDPs are, why they offer advantages over older alternatives, and what to look for in a CDP vendor.

What's a CDP?

Raab Associates defines a Customer Data Platform as a "marketer-controlled system that uses persistent, cross-channel customer data to support external marketing execution". Each part of the definition distinguishes the CDP from alternative approaches to working with customer data:

- "marketer controlled" means that marketers run the system directly. This distinguishes the CDP from conventional data warehouses or marketing databases, which are built and maintained by either the company IT department or an external vendor. CDP deployments do require some technical knowledge and some cooperation from corporate IT, but marketing still retains control.

- "persistent, cross-channel data" distinguishes CDPs from personalization systems that work only with information immediately available, such as behavior during a Web session, or with information captured within a single system, such as call center conversations. The CDP builds a comprehensive long-term view of the customer – that is why it exists.

- "support external marketing execution" distinguishes the CDP from marketing "clouds" or suites that are share data primarily among their own components. In theory, this sharing eliminates the need to create a separate customer database or to build custom integrations between that database and other systems. Unfortunately, many of today's clouds and suites have been assembled from purchased components and lack the degree of prebuilt integration marketers that really need. Nor do they fully expose their customer data to external systems. CDPs are designed to do exactly this.

CDPs are a relatively new type of system. The first generation started appearing about ten years ago. These firms built the cross-channel database primarily to support other functions offered by the vendor, such as predictive modeling or campaign management. This was necessary because the lack of such databases was preventing vendors from selling their tools. In the past few years, "pure" CDPs have emerged that

provide the database alone. The demand for such systems has grown as marketers recognize that such databases can support many different applications and can serve as a central coordination point for managing a unified customer experience.

Many of the newer systems extend this coordination to include some decision functions. These allow users to define rules that select single records or lists of customers that meet specified criteria. The records may be sent to other systems for analysis, message delivery, additional processing, or other actions. This approach adds value to the CDP without blocking marketers from using the analytical and execution systems of their choice.

Recent years have also seen a growth in real-time capabilities among CDPs. Older CDPs generally worked with batch data inputs. This was adequate when their primarily goal was to provide batch outputs such as lists of model scores or campaign segments. But as CDPs evolved to support a broader range of tasks, the need for real-time updates and access became more pressing. The newer CDPs, being more directly focused on their role as the primary corporate source of customer data, are more likely to provide real-time services.

CDP Benefits

The fundamental benefit of a real-time CDP is that companies get the unified customer view they need to create a superior customer experience across all stages of the life cycle: acquisition, growth, retention, and service. Specific benefits of building that view with a CDP instead of other techniques include:

- Faster deployment. CDPs are designed for quick deployment, using prebuilt templates and structures whenever possible to minimize work effort. The CDP project is much more limited in scope than a full data warehouse or cloud/suite implementation, again enabling quicker results.

- Lower cost. The same features that save time with CDP deployment also save labor. The CDP automates many tasks that otherwise require skilled technical staff, again reducing cost both during deployment and on-going operation.

- Greater marketing control. Marketers control the CDP directly, with minimal reliance on internal or external technical staff and without the constraints imposed by a prebuilt cloud or suite architecture. This control extends to letting marketers make many of the post-deployment changes, such as adding new data sources, processing steps, or execution systems.

- Easier growth. The ability for marketers to makes many changes on their own makes it easier for the company to adapt as new systems and channels are

implemented. The same features that make technical staff more productive during deployment also minimize technical effort for future adjustments.

- No vendor lock-in. CDPs are designed from the start to work with other vendors' systems, so they have no explicit or hidden obstacles to using whatever products the marketer prefers. This freedom lets marketers pick the right systems for their needs, including new needs that suite and cloud vendors can't always move quickly enough to support.

How They Work

In some ways, CDPs sound too good to be true: a custom-built, open-to-anything data warehouse that doesn't take years of specialized IT effort to build. But they're not snake oil and they're not magic. Rather, they apply a collection of new technologies that, taken together, change the boundaries of what can be done with customer data. Key factors include:

- Data capture. Collecting data from external systems has traditionally required a great deal of custom programming to define input formats, set up connections, transfer the data, and verify its accuracy. CDPs avoid much of this by providing prebuilt connectors and forms to collect the necessary information and credentials. These automatically generate the actual program scripts that would previously have been written by hand. Some CDPs avoid data transfers altogether by placing their own data-gathering tags on Web pages, email, display ads, and mobile apps. Some IT cooperation is still required to place the tags or get data access credentials. But the actual IT work is minimal and almost no further effort is required after the initial deployment. This is very different from a system where IT staff is involved continuously.

- Data storage. Conventional databases are like a particularly dense file clerk who must be told in advance exactly where to put each item and must be retrained to handle the slightest change. Most CDPs use unconventional data stores that effortlessly accept new inputs. This speeds the initial connections, eases subsequent changes, and minimizes the need for IT support. Because conventional tools often cannot access these unconventional data stores directly, many CDPs make it easy to automatically extract selected data into conventional databases.

- Identity association. This is the task of linking data from different systems that refers to the same customer. It's a challenge because the different systems often use different identifiers: email address, postal address, cookie ID, device ID, customer account number, social media handle, etc. There's no single new technology that gives CDPs an advantage in doing association. Instead, some CDPs include extensive applications of standard matching technologies while others rely on external systems. Because of this variation, marketers need to assess this

particular functionality in detail to ensure they select a CDP that fits their particular requirements.

- Data enhancement. This involves several different functions, including standardizing and validating addresses to help with matching; preparing data for analysis by calculating trends, aggregates, and model scores; and adding useful information from outside sources such as demographic or intent databases. Many CDPs do the enhancement automatically or allow marketers to easily choose which enhancements to apply. They take advantage of prebuilt connections to extensive enhancement data sets collected by third parties, the vendors themselves, or both.

- Data access. CDPs are expressly designed to allow access by external systems, so they have built-in features that had to be custom-created in older customer databases. These features allow access through modern methods, such as API calls, that are much easier to set up and maintain than older access techniques. At the same time, CDPs still support the older techniques, such as database queries and file extracts, to accommodate systems that can only use them. The CDPs let non-technical users set up access through both old and new techniques with minimal IT support.

What to Look For

While Customer Data Platforms are designed for non-technical users, picking the right one still requires understanding how its technology matches your technical requirements. If your marketing team lacks adequate knowledge to make these judgments, be sure to call in experts who can. Here are some key questions to ask:

- Data capture. The key issue is how hard it will be to connect with your source systems, both current and future. This set-up should be as painless as possible. You'll also want to make sure the system can capture all the different data types you'll need for your marketing projects. Specific questions include:

    - What types of systems can the system gather data from (Web, email, call center, mobile apps, social, etc.)?
    - What data types can it accept (structured records, freeform text, tagged text such as JSON or XML, video, etc.)
    - What types of feeds are supported (batch files, database queries, real-time)?
    - What types of APIs are available and what API functions are supported?
    - What source systems have prebuilt connectors available?
    - What steps are required to set up a new feed?
    - What steps are required to accept new data elements within an existing feed?
    - What data quality features are available to check data as it is added?
    - How quickly does new data become available?

  - Are there any limits on input volume?

- Data storage. You want to be sure the system can store all your data, both in terms of volume and variety. The goal is to keep the raw details whenever possible, because you never know exactly what will be needed in the future. Specific questions include:

  - How is data stored (relational database, NoSQL database, etc.)?
  - What data types are supported?
  - Are there any limits on storage volume?
  - Are there any constraints on the data model (number of objects, relationships among objects, attributes per object)?
  - Does the system store raw details or only summary records?
  - Are old values of standard attributes stored or simply overwritten (e.g., previous address)?
  - Are all data changes time/date stamped?

- Identity association. This is the heart of the complete customer view. Some systems have extensive identify association functions built in, while others rely on external services to do the association. External services are acceptable so long as the integration is easy and results are returned quickly.. Specific questions include:

  - Can the system associate records from different systems using a common identifier (e.g. email address, account number, device ID)?
  - Can the system associate records based on similarity (e.g. minor differences in name and address)?
  - Can the system associate records based on patterns (e.g., match on any 3 of 5 name/addrss elements)?
  - Can the system associate records based on activity (e.g., log into same account on different devices, send email from same address on different devices)
  - Can the system associate records based on probabilistic matching (e.g. simultaneous use of two devices in same locations over time)
  - Can the system associate and enhance records using external services (e.g. LiveRamp, Tapad, Neustar, etc.)?
  - What external services have existing connectors to the system?
  - Can the system set and manage identifiers such as cookies
  - Can the system track anonymous users and retain their history when they are identified
  - Can the system support multiple layers of association e.g. individual vs household, employee vs business

  - Can the system create a 'golden record' with the best known value for each data element associated with an individual
  - Can the system apply restrictions or permissions such as email opt-in or telephone opt-out across all related records?

- Enhancement. The systems should be able to clean up your data to make it usable. This is a key step before some types of identity association and can also create summaries that save time during queries by external systems. The system should also be able to call on external resources to add new data and to help with standardization and validation processes. Specific questions include:
  .
  - Can the system standardize inputs such as names, addresses, and titles?
  - Can the system verify data elements based on format (phone number, email address, etc.)
  - Can the system verify data elements using external sources (postal files, email services, Web domains)?
  - Can the system create derived values such as aggregates, trends, time-since, and model scores?
  - Can the system add data from external sources such as geography-based cluster codes, demographics, intent databases, audience memberships, etc.
  - Does the system track the date, source, and any usage restrictions on enhancement data?
  - What sort of information does the system provide about the quality of data it contains?

- Data access. All data in the system should be easily accessible to external systems. Speed is a primary concern for real-time interactions. Prebuilt connectors are extremely important, as is a powerful and well-documented API. Specific questions include:

  - What steps are required to expose data to external systems?
  - How can external systems access the data (SQL queries, file exports, API calls, etc.)?
  - Can external systems access all data or only selected elements?
  - Can external systems access detail data or only summaries?
  - Can the system push messages to external systems in response to user-defined trigger conditions?
  - Can external systems request data about a single individual?
  - Can the system find data about an individual based on incomplete information (e.g. fuzzy match) or does it require a known identifier (e.g. account ID or email address)?
  - What data is available for queries about a single individual?
  - How quickly can the system return data about a single individual?

- What systems have prebuilt connectors to access the data?
- Are there any limits to the amount of data that can be read or extracted in a single query?
- Are there any limits to the amount of data that can be read or extracted within a given time period?

Summary

Customer Data Platforms fundamentally change the process of building a complete customer view. They offer marketers an approach that is faster, cheaper, more effective, and more flexible than previous alternatives. Although it will take some effort to evaluate these new options and pick the system that's best for your company's needs, the advantages are worth the effort.

About Raab Associates Inc.

Raab Associates Inc. is a consultancy specializing in marketing technology and analytics. Typical engagements include marketing architecture planning, business needs assessment, technology audits, vendor selection, results analysis, and dashboard development. The company also consults with industry vendors on products and marketing strategy

Contact:
Raab Associates Inc.
730 Yale Avenue
Swarthmore, PA 19081
www.raabassociatesinc.com
info@raabassociatesinc.com

