

# Customer Data Platforms for Enterprise Marketers



**Published by:**



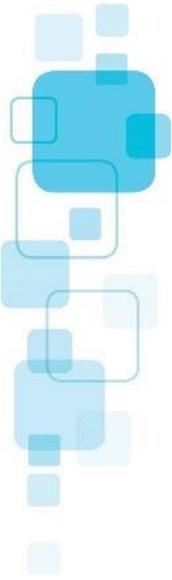
**Sponsored by:**



4235 South Stream Blvd, Suite 410

Charlotte, NC 28217

[www.quaero.com](http://www.quaero.com)



## Introduction

Customer data has become central to marketing, as customers demand personalized experiences across all channels. In addition, insights from customer data impact customer service, corporate strategy, and digital transformation efforts aimed at personalizing customer experience at scale. A recent McKinsey study<sup>1</sup> found that companies that harness creativity and data in tandem have growth rates twice as high as companies that don't. Yet an Experian study<sup>2</sup> found that lack of data and insight is the single biggest challenge organization face when it comes to digital transformation.



Enterprise marketers face even greater challenges than their colleagues at smaller companies. These include larger and more varied data sets; complex technology infrastructures often including a mix of new and old systems; global customer bases subject to different market and regulatory conditions; massive organizational structures; and powerful central IT organizations that often focus more on security and reliability than meeting customer expectations.

But enterprise marketers also have more resources. They can often draw on large teams of analysts and data scientists; sophisticated technology staff including marketing technology specialists; highly trained internal and external marketing experts; a wealth of opportunities to run tests; and budgets to build or buy advanced systems; and the scale to justify the cost of advanced systems.

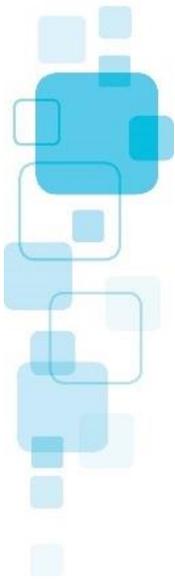
One result of all these factors is that enterprise marketers tend to buy single-purpose “best of breed” products rather than settle for multi-function packages that cost less and do many things but are not equally good at all of them. While this makes business sense, it leaves the enterprise marketers facing the task of integrating data from many separate products. The problem remains even at enterprises that have purchased “marketing clouds” from companies like Adobe, Salesforce, or Oracle. Those “clouds” are combinations of systems that the cloud vendors have purchased. But while the cloud components are sold under the same logo, they are barely more integrated than they were when they were sold independently. So even enterprises that commit to cloud systems find themselves faced with the challenge of knitting together data from separate products.

This, then, is the enterprise marketing data dilemma: best-of-breed systems are individually the best solution for each function but collectively create fragmented data that prevents enterprise marketers

---

<sup>1</sup> The most perfect union: Unlocking the next wave of growth by unifying creativity and analytics. Brian Gregg, Jason Heller, Jesko Perrey, and Jenny Tsai, McKinsey & Company, 2018

<sup>2</sup> Mastering the customer experience: Why insights-driven businesses are winning in the digital age. Experian, July 2018

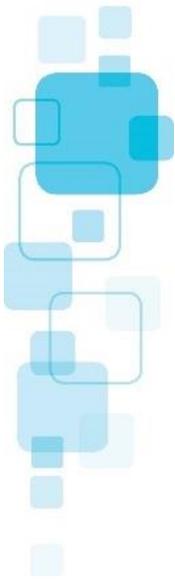


from delivering the integrated experiences their customers expect. This fragmented data environment also makes it more difficult to understand customer life time value and develop broader customer strategy. This paper explores how the right Customer Data Platforms can resolve this dilemma so that marketers, IT departments, and customers all come out as winners.

## Solving the Integration Dilemma

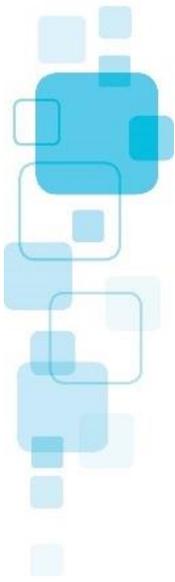
Before discussing CDPs in depth, let's start by defining the underlying requirements. What, exactly, does a successful customer data solution need to provide? Capabilities include:

- **Gather data from all sources.** Any major enterprise has dozens – sometimes hundreds – of systems capturing customer data. These include primary customer facing systems such as order processing, point of sale, and customer service, and specialized systems such as retail kiosks, return tracking, and warranty registration. Marketing alone has data coming from Web sites, mobile apps, email, display advertising, paid search, social media, content distribution, field events, and more. This data comes in structured formats such as purchase records, semi-structured formats such as Web site logs, and unstructured formats such as call center notes and social media comments. Newer formats appear regularly, including voice, video, digital photos, and text messages. Even more are on the horizon: augmented and virtual reality interactions, location tracks, emotional status, and data streams from connected devices ranging from fitness trackers to home security systems to smart appliances. All of these need to be gathered and stored efficiently, often in real time. Although marketers will inevitably need to prioritize which sources they capture first and how long they keep the data, there's a strong tendency is to ingest and retain as much as possible because so much of it has a clear potential to be useful.
- **Retain full detail.** When data processing costs were high, marketers needed to filter and summarize their data as soon as possible. Today's "big data" technologies make it easier and cheaper to store the original details and then summarize later when a specific purpose emerges. This may still require some transformations to make the stored data more easily accessible by putting it in standard formats and creating access paths such as indexes. But the ability to reconstruct the original detail is still essential for many applications such as journey analysis and training machine learning systems.
- **Extract meaning.** Even as the raw data is being stored, the system must also start to extract meaning such as identifying the customers, products, topics, locations, media, and other items that will later be the subjects of analysis. As volumes get larger, it becomes more important to do this as part of the initial processing flow rather than storing the data first and then retrieving it for a second look. Unstructured and semi-structured data, as well as non-text formats like speech and video, take especially sophisticated processing. This conversion is essential for most analytical and customer management processes, which can only work efficiently with structured data.
- **Create unified profiles.** Once the data is ingested and placed in a usable structure, the system needs to assemble customer profiles by combining information that relates to the same individual. This is a highly complex task that can employ different matching methods for different purposes. At a minimum, any system needs to bring together records that have the same identifier, such as a customer number. In addition, any system should be able to stitch together records based on overlapping identifiers: such as a phone number on one system linked to an email address from another system via an account record that contains both the phone number and email address. Done correctly, the identity management system will create a persistent ID that retains these



connections over time even if any single identifier changes. Other matching methods rely on data patterns to link records where there is no exact match between a unique identifier: for example, a system might look at name, street address, birthdate, and zip code and create a match if there's a match or near-match on any three of the four. Still other methods look at behavioral data such as linking two mobile devices if they appear frequently in the same locations at the same time. Nearly all matching methods may draw on external systems that use additional data sources to find matches that would be missed looking at internal data alone. Once these linkages are established, all the data related to the identifiers should be pulled together to create a full profile of each individual.

- **Maintain data security and personal privacy.** Customer data is increasingly subject to regulations that control how it can be collected, protected, and used. Having a single system that acts as central source of customer data can greatly simplify compliance but also means that system itself needs to meet all the various requirements. This means it must track the origins and permissions associated with each data element, enforce constraints on who can access it for which purposes, and keep records to prove the rules have been followed. The system also needs to incorporate security processes and techniques, such as encryption, needed to guard against unauthorized use. In many cases, the system will need to be certified against external standards. Beyond legal requirements, many companies will have internal policies that further govern data privacy and protection.
- **Structure data for use.** After all data for each customer is connected into a single profile, additional processing is needed to make the data available for efficient use. This may include standardizing data from different sources, identifying events, creating aggregates, adding derived variables, calculating model scores, assigning segment codes, and creating indexes to speed access. In some cases, the data may be placed in entirely separate structures such as a flat file for use by predictive modeling systems or star schema data cubes needed by business intelligence tools. The system may be required to generate specific outputs on a regular basis, such as daily files for reporting systems or “golden records” with the best version of each customer data. Although some technical skill will always be needed, it should be possible for non-technical users to perform the bulk of these tasks for themselves. This allows the system to quickly adjust to new needs as they arise.
- **Allow easy external access.** The system needs connectors that enables external systems to read its data. These may involve direct queries against the main data store, access to derived files in special formats, or API calls that extract subsets of the data. There may be specific performance requirements where real-time response is needed and there may be need for on-demand processing such as updates to model scores based on events during a current interaction. The system needs to provide a well-documented access method and simple connections, ideally requiring little more than a password and file destination. Prebuilt connectors to specific systems are important for smoothing deployment and to access system functions, such as matching against a new identifier, in addition to the data itself. Security and controls to avoid unauthorized use are especially important when governing data access.



- **Easily adapt to change.** Beyond these specific capabilities, the customer data system must easily adapt to new data sources, processes, and clients. This is essential to keep up with the pace of change in marketing systems and the larger corporate environment. In particular, it should be possible to add a new data source without explicitly mapping all data elements and values into a structured data model, to access loaded data without specifying the available elements in advance, and to define output formats and destinations without advanced technical skills. Similarly, the system should be able to automatically add data storage and processing resources as needed to achieve acceptable performance.

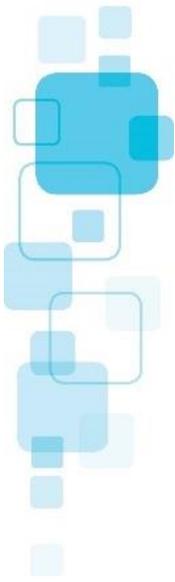
### Existing Systems

Most enterprises already store customer data in systems such as data warehouses, data lakes, Data Management Platforms, Customer Relationship Management, and marketing automation. These are important sources of customer information but generally cannot provide a unified customer database. See the Appendix for a more detailed discussion.

## Customer Data Platforms

Customer Data Platforms are a relatively new class of system designed specifically to build and share a unified customer database. Because this is their original goal, they generally meet the requirements listed in the previous sections. This offers many advantages for enterprises that are straining to make the best use of their customer information. Advantages include:

- **Design focus.** CDP vendors are free to focus on the single challenge of building a unified, accessible customer database. This lets them select appropriate technologies and data structures. Equally important, they can build processes to simplify system deployment and ensure efficient operations. Because building the unified database is a very complicated task with many components, the freedom to work on these without also meeting other requirements is extremely important.
- **Source and destination neutral.** CDPs are designed to connect with any data source and to share data with any destination system. This avoids the limits of unified customer data systems that are part of a single-vendor marketing suite and connect best with other products from that same vendor. Neutrality comes most easily to independent CDP vendors, who have no financial incentive to favor connections with any other vendor's systems. But it's also technically possible for a marketing suite or marketing cloud vendor to create a connection-neutral CDP. It's up to buyers to find out whether they have really done so.
- **Advanced ingestion and identity management.** These are the two technical requirements that most distinguish a CDP from other systems. The specific requirement for ingestion is that adding new sources be as easy as possible: a CDP should accommodate new inputs with little or no manual data mapping. CDP vendors vary in how well they do this but it's a primary consideration for all of them, unlike other types of systems. Similarly, CDP vendors vary in the details of their identity management but all provide fundamental stitching and persistence capabilities that other types of systems usually lack.
- **Reduced demands on technical staff.** CDPs are packaged software. This means they include predefined process flows for key functions and interfaces to control them. These give a starting point in getting the system set up and make managing the system easier than working directly with the underlying technology. The result is faster deployment, less development risk, and more complete functionality than a custom-built system. This does not mean that CDPs can be deployed



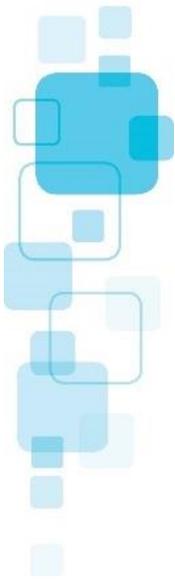
with no involvement from IT: the corporate IT group controls access to customer data and are experts at what's in the company's own systems. Their contribution remains vital even as the CDP reduces the amount of effort required from them.

- **Avoid vendor lock-in.** The CDP provides a central repository for customer data that is separate from the systems that generate it. This means buyers can change source or destination systems without fear of losing whatever data they contain. This flexibility is especially important to enterprise marketers, who are most likely to want the best possible product for each purpose and thus to switch vendors when a superior new product becomes available. It also reduces dependence on marketing cloud or suite vendors, even if the company buys most of its systems from one of them. This benefit persists even if the CDP itself is provided by the suite or cloud vendor, so long as that vendor's CDP is truly open to other vendors' products.
- **Enable orchestration.** Many enterprises are working to orchestrate customer treatments across channels, even when interactions on those channels are run by separate systems. Unified customer data helps with orchestration because all channel systems are able to work from the same data. Companies that take the next step and install a single orchestration engine to drive interactions across different channels get the further benefit of being able to connect the orchestration system to a single data source. Without a CDP, the orchestration engine would likely need to connect to multiple data sources and effectively build its own unified customer database. Avoiding this extra effort simplifies deployment of the orchestration system and lets users pick the best orchestration engine without concern for its database-building features.

## What to Look For

We've seen what companies need in a unified customer database and why a CDP is often the best way to meet those needs. But all CDPs are not alike. In fact, there's probably more variation among CDPs than in other system categories, largely because vendors from different backgrounds have recognized the need for CDPs and converged on the market from several directions. Adding even more confusion, marketing suite and cloud vendors have recognized the same need and added CDP-like functions to their systems, whether or not they deliver a fully qualified CDP product. As you sort through the options, here are considerations to help find the right choice.

- **Requirements.** The requirements listed above are the bedrock of any CDP. But every system calling itself a CDP won't necessarily meet them and systems that do meet them can vary in the details. In addition, each company has its own needs, so you'll need to develop a modified requirements list based on your own situation. Once you've gone through that process, you'll have a checklist that you can use start qualifying potential CDP systems.
- **Right scope.** Many CDP vendors provide predictive and personalization functions in addition to building the core database. This can save integration effort and reduce costs but it's only useful if you want the vendor's version of those tools. Enterprise marketers and data scientists often prefer best-of-breed tools that they've already put in place. If that's your situation you probably don't want to pay for features you won't use. Then again, the CDP's version of these tools is sometimes more automated or easier to use, in which case it can complement rather than replace existing products.
- **Support for existing systems and channels.** Prebuilt connectors are one of the ways that CDPs speed deployment and reduce costs. Some vendors have dozens of prebuilt connectors; others have

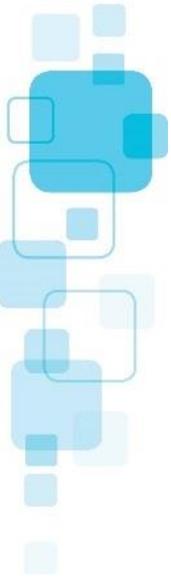


just a handful. Of course, what matters more than the total count is support for the specific systems your company uses or expects to use. As you consider CDPs connections you will also want to assess their ability to connect to your custom systems and configurations. Similarly, some channels have certain features or requirements that are common to all systems in those channels: again, you'll want to find a CDP that supports the channels you use.

- **APIs to connect new systems and channels.** Nearly all CDPs provide Application Program Interfaces (APIs) that let them connect to new systems if no prebuilt connector exists. If it looks like you're going to need new connections, you'll want a system with APIs for this purpose. Be sure to examine the APIs to ensure they are well documented and have the functions you'll need.
- **Scalability.** Scalability has many dimensions, including number of data sources, types, objects, and attributes; number of customers; total data volume; peak transaction volumes; required response times; allowable batch update windows, latency between data receipt and availability, and more. You'll need to identify the likely bottlenecks in your particular situation and explore those carefully. This will often require running tests on the actual system, especially if you are working at larger scale than the vendor has already supported successfully.
- **Deployment and support model.** CDPs vary in their flexibility of deployment. Some deploy only on specific public cloud environments (e.g. AWS or Azure). Others can run on public clouds, private clouds, or a corporate data center. Some CDPs are pure software companies that partner with service providers to provide comprehensive support and service, others offer a full range of professional and/or managed services. The best choice may depend on your internal capabilities as well as your relationship with your IT and infrastructure teams.
- **Security.** This also has multiple dimensions, including controls over external data access, controls over internal data access and functions, protection against hardware and software failures, and compliance with privacy regulations. Enterprise systems often must support large numbers of users with different levels of authority and access to different subsets of data based on geography, department, job role, and other variables. The CDP should have internal controls to restrict access and adjust its interface so users are only presented with information and options they are authorized to use. These functions should be designed so they can efficiently manage large numbers of users. They should also integrate with other corporate systems that manage user identities and authorities.
- **Vendor.** As with any vendor relationship, you'll want to consider the vendor's experience, stability, resources, reputation, pricing, support, and other considerations.

## Conclusion

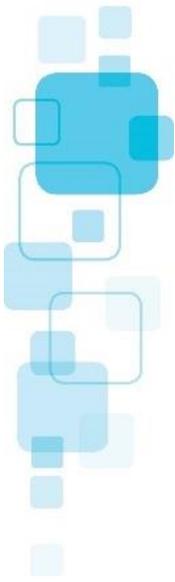
Unified, accessible customer data is essential for today's enterprise. Yet even the largest and most sophisticated organizations often struggle to get the data they need. One reason has been the belief that existing systems can, or at least should, be adequate. But existing systems were built for other purposes and in most cases cannot successfully support the special requirements of a unified customer database. Customer Data Platforms are a new type of system designed especially for this purpose. In many cases, CDPs will offer a more effective solution, faster and at lower cost, than existing alternatives. Although selecting a CDP takes careful research, the result should be a system that gives businesses the data they need to give customers the experiences they expect. It is well worth the trouble.



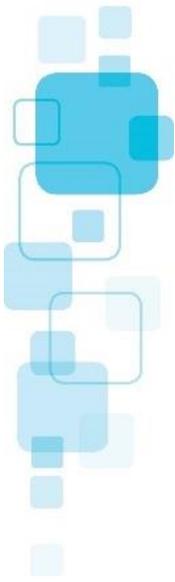
## Appendix: Other Customer Data Systems

Enterprises have been working with customer data for many years. Most will have several types of customer data systems in place. It makes sense to ask whether any of those existing systems can meet the requirements defined above. In most cases, the answer will be no. The broad reason is the existing systems were built for specific purposes and cannot easily be adapted to act as a unified customer database. The specific reasons are different for each system. Here's a quick look.

- **Data warehouse.** Most enterprises will have at least one data warehouse and often several. Data warehouses are built to combine data from multiple sources and make it easily available for analysis. They are also built to serve the entire enterprise, with marketing being one of many stakeholders. The primary focus on analysis leads data warehouse designers to use relational databases and limit their systems to structured data sources. They then summarize the data and place it into special relational structures (“star” or “snowflake” schemas) that allow efficient analytical queries. This is an impressive technical achievement (relational databases really weren't built for analytics) but it means that the databases must be carefully designed, which in turn means adding new data elements takes careful thought and considerable technical skill. Placing the input data into the warehouse data structures can also take many hours of processing as the system builds aggregates and indexes needed for quick performance. This is another reason that any changes must be carefully considered. Even when this work is done, a standard data warehouse doesn't support real-time access to all details about a single customer, as often needed for customer interactions. Warehouse developers sometimes address this by creating an extract file (“operational data store”) that makes a pre-specified subset of data available in real time. But even this approach is limited by reliance on the data warehouse as the original source: because the data warehouse is largely limited to structured data as inputs, it cannot provide the complete customer view that includes unstructured and semi-structured data as well.
- **Data lake.** The rigidities of data warehouses have led many enterprises to supplement them with data lakes, which are in some ways the polar opposite: repositories that accept any kind of data with no preparation. This avoids the painstaking data preparation required to load a warehouse – or, more accurately, defers that preparation until a later step in the process. But it does provide data scientists and business analysts with easy access to all the ingested data, saving them the often-substantial labor of tracking it down and extracting it from the original source systems. While a data lake might be a source for a unified customer database, it doesn't include all the subsequent processing needed to turn the data lake contents into usable customer profiles.
- **Data Management Platform.** DMP's are used by marketing departments to manage online advertising audiences. They contain individual-level customer profiles, although these are often based on cookies or devices that are not tied to identified individuals. While an enterprise's internal DMP may house a combination of known and anonymous identifiers, external DMPs often contain only anonymous information to avoid privacy issues. The primary design goal of DMPs is return audience data quickly, either to generate lists of profiles that meet specified criteria or to decide what to bid on the opportunity to serve an ad to a single profile. This limits each profile to a string of summary attributes, such as gender, region, interests, products purchased, and in-market status. While there may be several thousand of these on a profile, they still do not capture the full detail of Web browser logs or transaction details. So, the DMP provides a much more limited customer profile than required for a unified customer database. DMPs also have limited or no built-in identity matching.



- **Customer Relationship Management.** CRM systems are often an enterprise's primary store for customer information. They typically gather information directly through customer interactions managed within the system by sales and customer support teams. To do this efficiently, they use relational database structures that are optimized for fast access and updates to individual customer profiles but are not good analyzing or extracting data about large groups of customers. This also limits their ability to handle unstructured or semi-structured data, apart from sometimes storing it in large blocks. Because they work primarily with data that users have entered directly into the system, CRMs have limited abilities to ingest external data and to do advanced identity matching.
- **Marketing automation.** Marketing automation systems are used by marketing departments to manage customer and prospect relationships. Like CRM systems, they are designed primarily to work with data they have generated internally, such as email histories. Some do provide specialized connections to external sources such as tags to capture Web site behaviors or synchronization with CRM systems. Most use a relational database and are largely limited to structured data, although a few top-end marketing automation products have added support for semi-structured and unstructured sources. Again, like CRM systems, they have very limited data quality and identity matching functions and offer limited support for external data access and analytics.
- **Analytics environment.** Many companies have created a stand-alone analytics environment that pulls data from a data warehouse and/or multiple systems and connects it to a workbench with analyst tools to transform it, build predictive models, do pricing analysis, create strategic segmentations, and calculate lifetime value. These environments often include a snapshot of customer data or limited time series. This approach often creates challenges in terms of not having a single source of truth, integrating model scores into execution systems, and measuring model performance after campaigns.
- **Integration platforms.** Integration platforms are a relatively new class of system that is designed to connect other systems, primarily to support cross-system business processes. They do this by mapping data from different systems so they can transport it from one system to another. Because their role is only to move data between systems, they do not store it internally as a customer profile. This prevents them from matching customer identifiers over time, building customer histories, tracking trends, and performing other tasks needed for a unified customer database. Where they do overlap is providing multiple systems with a consistent, consolidated view of customer data but with the crucial limit that they can only share current information in the source systems.
- **Marketing Service Providers.** Although this approach is becoming obsolete, many companies still depend on external agencies or specialized marketing service providers to build and maintain their customer data. These solutions are convenient for marketers in that they let them sidestep their IT organizations, but they reduce flexibility and agility in an age when customer data is increasingly recognized as a strategic and competitive asset.



## About Quaero

Quaero accelerates data driven teams as they transform disconnected customer data into energized assets. We have packaged nearly two decades of experience into Quaero's Customer Data Platform (CDP) delivering a turnkey solution for monetizing customer data from any source - digital or offline. We deliver persistent 360° views of customer data to support key use cases for clients across multiple industries including Retail, Consumer Products, Financial Services, Media, and Health Care.

Quaero is purpose-built for end users who live and breathe data daily. Marketers and Data-As-A-Product teams use our platform to source and activate information from across their enterprise, turning their unique customer knowledge into bottom line outcomes. We release data scientists, analysts, and engineers from time-consuming data cleansing chores, enabling them to focus on delivering powerful segmentations, predictive models, and concrete additions to the company's data balance sheet.

The Quaero CDP enables fully automated data management unified by a state-of-the-art identity resolution engine. Standard connectors, no-code configurations, and privacy-by-design ensure functional and technical team members work jointly in a seamless, secure customer data environment rapidly deployed in your preferred cloud or internal infrastructure. Quaero's Customer Data Platform compresses the time, cost, and risk of building strategic assets while keeping fast moving teams compliant and in firm control of their proprietary data resources.

### Contact:

Quaero  
4235 South Stream Blvd, Suite 410  
Charlotte, NC 28217  
[www.quaero.com](http://www.quaero.com)  
[info@quaero.com](mailto:info@quaero.com)

## About the CDP Institute

The Customer Data Platform Institute educates marketers and marketing technologists about customer data management. The mission of the Institute is to provide vendor-neutral information about issues, methods, and technologies for creating unified, persistent customer databases. Activities include publishing of educational materials, news about industry developments, creation of best practice guides and benchmarks, a directory of industry vendors, and consulting on related issues.

The Institute is focused on Customer Data Platforms, defined as "packaged software that maintains a unified, persistent customer database which is accessible to external systems."

The Institute is managed by Raab Associates Inc., a consultancy specializing in marketing technology and analysis. Raab Associates defined Customer Data Platforms as a category by Raab Associates in 2013. Funding is provided by a consortium of CDP vendors.

For more information, visit [www.cdpinstitutue.org](http://www.cdpinstitutue.org).