Enterprise-Class Businesses Can't Afford Bad Data

Poor data quality and limited data access are costly. Only an enterprise-class dynamic customer data management platform will maintain the data quality needed for today's real-time world.
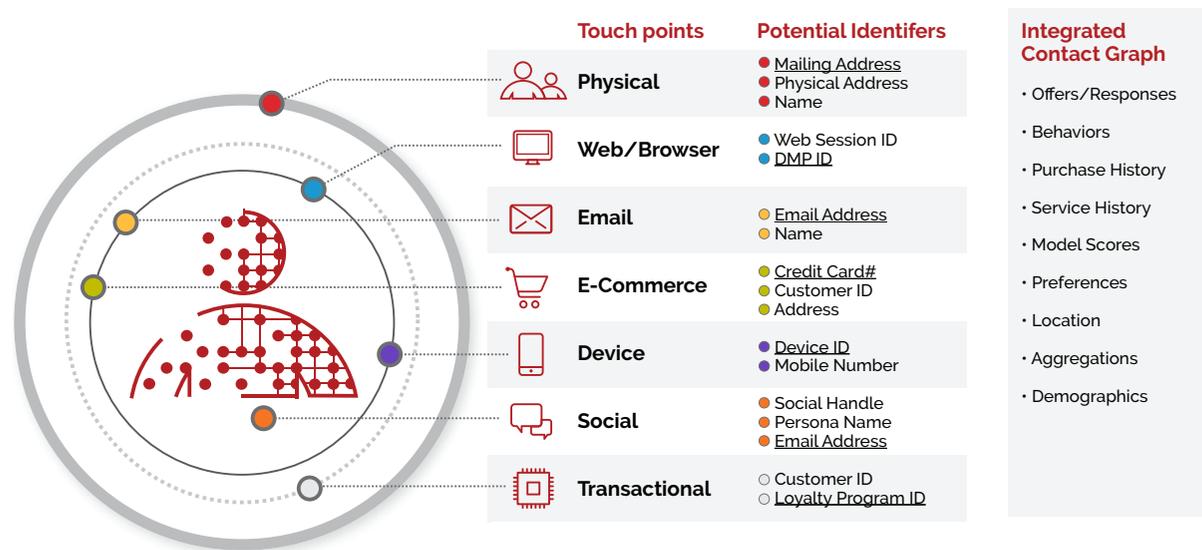
redpoint™

Table of Contents

Most marketers and data scientists know that bad data is costly. But consider this stunning figure: The yearly cost of poor-quality data is $3.1 trillion, in the U.S. alone, according to IBM estimates.

That's why marketers who need easy access to large data streams may need to rethink their current data practices. Many marketers, for instance, flow those large streams of data into a data lake. All too often the lake is stagnant and data quality suffers. Moving at the cadence of the customer —a must for marketers at large enterprises aiming to meet customers' expectations for a relevant experience—requires access to a dynamic data repository that continuously refreshes: **a robust, dynamic customer data management platform (CDP)**.

Not just any CDP will do. The platform must be enterprise-class:

- Able to **ingest** data from internal and external sources in batches and in **real time**
- Provide **quality** assurance (e.g., merging, **matching**, and **identity resolution**)
- Create an **identity graph** of each customer along with all profile and transactional history (i.e., a **Golden Record**)
- Offer automations and workflows that let marketers focus on strategy while the system handles **personalized interactions at scale**

Redpoint Golden Record = Identity Graph + Contact Graph

Touch points	Potential Identifers
Physical	Mailing Address / Physical Address / Name
Web/Browser	Web Session ID / DMP ID
Email	Email Address / Name
E-Commerce	Credit Card# / Customer ID / Address
Device	Device ID / Mobile Number
Social	Social Handle / Persona Name / Email Address
Transactional	Customer ID / Loyalty Program ID

Integrated Contact Graph

- Offers/Responses
- Behaviors
- Purchase History
- Service History
- Model Scores
- Preferences
- Location
- Aggregations
- Demographics

The contact graph represents a unique individual, who may exist in fragmented forms across many touchpoint, each with potentially different identifiers and related data assets.

Data pulled from a "no-frills" CDP or stagnant data lake (aka, data swamp) may be fine for a small business, but it creates poor customer experiences for large organizations. Customers suffer confusion and a barrage of noise due to irrelevant campaigns and communication. Data streamed from a comprehensive, dynamic CDP enables marketers to engage customers instead of exasperating them.

When you're trying to gather customer insight that will help you deliver a profit-generating customer experience (i.e., so relevant that customers feel valued and demonstrate their appreciation by spending more, more often), it's essential to understand the differences between a basic CDP or data lake and a robust, dynamic one—as well as the role martech plays in data management. It's also important to understand how the technology evolved and where it's headed, so you can get the most from it.

Data Warehouse	Summary-level data
Data Lake	Highly variable data - varying consistency & accuracy
Dynamic Enterprise CDP	Highly accurate & consistent data

Data streamed from a comprehensive, dynamic CDP enables marketers to **engage customers instead of exasperating them.**

redpoint™

State of disarray

In the heyday of the emergence of big data, countless pundits declared and expected that the world was going to change forever because of Hadoop and other NO-SQL alternatives. Others asserted that all businesses were going to be run by code; code would do everything, and everyone needed to become a coder.

Much of this was a reaction to the world of data warehousing, where granular data was destroyed to create summaries for reporting purposes. Data scientists led the "big data march" for granular atomic data and quickly became the new bottlenecks to data access and usability; for example, they have to build the high-value models *and* the simple marketing-friction type models such as "customer lifetime value" and "mailbox optimization time."

This legacy has persisted to some extent, but also ran into the obvious realities that most companies are not going to be run by coders. Even the most ardent supporters of "everyone should code," realize that purely code-based systems are never going to scale to the extent needed for true enterprise-wide adoption; code-based projects would remain as special projects and gap fillers.

Even the most ardent supporters of **"everyone should code,"** realize that purely code-based systems are never going to scale to the extent needed for true enterprise-wide adoption; code-based projects would remain as special projects and gap fillers.

This trend has taken hold in the world of analytics. Most of the analytics for AI functions such as machine learning are done using tools like Jupiter notebooks, Python scripting, and Spark scripting, versus using any kind of simplistic user interface. So, unfortunately, although AI has proven its efficacy in numerous aspects of business, it has not scaled the way anything with that level of business value typically would. The reason goes back to the trend of doing everything with code, which does not scale for operational and CX use cases.

When you do everything through code, everything you do is piecemeal.

For example, if you are going to process data, you have to write code to extract/transform/load (ETL) it to make it fit for purpose. ETL functions need to be easily adaptable and scalable as data is forever changing; instead, ETLs are often objects—code snippets that process data and bring it into a location or transform data in some way.

This code-centric approach has had an impact on marketing and martech, as well. Many analytic tools often closely associated with marketing are often code-heavy data environments that use a storage-layer and compute-layer model to capture and process data. This requires data to stream into a data lake and pool there, losing quality, turning the lake into a swamp. The data lake becomes a kind of catchall, where analysts should be able to go in and access their data, process and build their models, and then come up with some magical insight that will transform the company. But that's rarely or episodically the case.

The challenge in trying to build an open-ended marketing data source like a data lake is a potential lack of data quality and governance tools that ensure the data is maintained in a healthy, clean, current, continuously updated, and properly matched and harmonized state. This state of disarray is where most of the CDP industry lives right now.

When data environments, such as a data lake, are split into a storage layer and a database or compute layer, very little data quality assurance is applied to the data as it is being streamed into the data lake. And all the processing of the data into a form useful for marketing then has to be coded. The unfortunate result of this is that the marketer is left with poorly executed identity resolution at best, minimal behavioral information, and no transformations (e.g., a year-over-year change in spend). It is just basic raw data.

This introduces another problem. It pushes transformations downstream, so marketers need to build their own aggregate values, derived attributes and other critical information. Transformations are where you get the real value from data, so missing it or doing it wrong comes at a great cost. When this transformed data is left out of segmentation or predictive models, it crushes creativity and the effectiveness of marketing programs. It also may result in irrelevant messages when these are processed outside the cadence of the customer, or even worse when they are processed inconsistently. All of this downstream activity is sub-optimal.

These data gaps are rampant in all manner of CDP—from orchestration or segment development technologies to marketing clouds. For instance, a marketing cloud may enable users to build audiences, but may lack accuracy because the underlying data is so limited.

The other effect that data lakes create is the natural inertia that resists large amounts of daily transactional updates, because all that data has to be processed using code. One marketing cloud vendor recently complained because its customer—a large retailer—had something in the tens of gigabytes for daily updates and that was too much for the vendor. Conversely, one of Redpoint's retail customers can easily process petabytes of transactional data every few days using the rgOne platform, a dynamic enterprise-grade customer data management platform. So, I leave it to you to assume which of those two retailers has the better marketing data to work with based on daily updates of trends, responses, sales, customer loyalty data, and more.

Data lakes create **natural inertia that resists transactional updates**

Blocks ROI $\$$

rgOne delivers perfect marketing data based on frequent updates of trends, responses, sales, etc.

Perfect Path to Revenue and ROI   **$\$$**

Much of this was a reaction to the world of data warehousing, where granular data was destroyed to create summaries for reporting purposes. Data scientists led the "big data march" for granular atomic data and quickly became the new bottlenecks to data access and usability; for example, they have to build the high-value models and the simple marketing-friction type models such as "customer lifetime value" and "mailbox optimization time."

This legacy has persisted to some extent, but also ran into the obvious realities that most companies are not going to be run by coders. Even the most ardent supporters of "everyone should code," realize that purely code-based systems are never going to scale to the extent needed for true enterprise-wide adoption; code-based projects would remain as special projects and gap fillers.

Is your data store a pristine data stream or a swamp? Swamps lack these critical components:

Component	PRISTINE DATA STREAM	SWAMP
Pristine, relevant, and high-quality data	✅	🚫
Organization and curation	✅	🚫
Enterprise access – for both operations and analytics	✅	🚫
Data governance	✅	🚫
Automated data quality processes	✅	🚫
Streamlined and auditable data changes	✅	🚫
Flexibility to modify to fit new data or new use	✅	🚫

Closing the gap

Data lakes and their lack of quality and governance aren't the only issue. There's an intrinsic problem in the marketing cloud world: The platforms are composed of tools that were standalone packages and acquired and integrated over time. All these tools have different underlying data structures and require essentially a "crosswalk" table to enable somebody to match "Chris A. Customer" on one system with "chriscustomer@gmail.com" on another. This introduces the potential for errors, data latency, and complexity for the user across the marketing cloud products.

Another limitation is that users really only have a few channels to work with: the ones that are represented by those cobbled-together tools. For instance, you may miss call center or in-store or mobile channels, but those channels are rife with essential customer data.

There's also an issue with marketing-orchestration technology and segment-development technology. Each one is a niche technology focused on a particular area that supports marketing, but the marketer still has to do the personalization themselves, versus using a platform that does the work for them.

redpoint

Similarly, many of the client-side personalization tools that enable building customer profiles may use an enormous volume of web behavior to drive personalization but lack access to other data sources that provide a complete picture of a customer. As a result, users need to take a Lego approach and link building blocks of tools that provide access to other data. Some of the building blocks themselves may be simple to use, but you have to have people using each one of those blocks every day. This adds layers of complexity to the process. One large U.S. retailer, for example, could not scale their A/B tests as they had to build all of the audiences, aggregates, and summarizations by hand.

The Lego approach isn't inherently bad; it's just not scalable nor able to deliver repeatable results.

> **Sure, Legos are fun to play with, but you wouldn't build a real bridge using them, because outcomes matter with a bridge. Outcomes also matter to marketers, so they need correct, clean, consistent data along with the right aggregates and KPIs to get to revenue growth.**

When optimizing return on your marketing dollar, there has to be structured process and data environment to ensure that the everything in the

91% YOY Revenue Improvements? Yes, Thank You

Xanterra took a structured approach with rgOne to drive ROI

Complexity hasn't stopped Xanterra Travel Collection from getting the most from its data. The hospitality company owns or manages a wide-ranging portfolio of properties, including national park lodging, a boutique cruise line, resorts and hotels, and specialized adventure tour operators. Xanterra uses Redpoint Global's rgOne platform to consolidate data from all of its properties and systems and build rich 360-degree profiles for virtually every customer and prospect. Xanterra regularly analyzes that data to find unique customer engagement opportunities. For example, it was able to identify seven key customer segments, craft detailed personas and experience maps for each, and personalize communications at scale, which the company does throughout the entire customer lifecycle.

Using the rgOne platform, Xanterra can quickly plan and execute new initiatives aimed at cohorts such as seniors, families, couples, or holiday travelers; cross-sell to prospects interested in multiple properties or forms of travel; and build loyalty by enhancing multiple aspects of the guest experience.

The results have been outstanding, including 91% YOY revenue improvements and 101% YOY transaction increases. In one specific circumstance, by combining segmentation, A/B testing, and its 360-degree customer view, Xanterra developed highly relevant offers sent via email with unique and targeted imagery, content, and subject lines; so relevant, in fact, that the company earned 73 cents per email, compared to 8 cents for previous average campaigns—an 839% improvement.

marketer's control—e.g., spend, messages, audience selection, channel selection, sequences—is being done correctly to realize the best possible yield.

Plus, think of the expense of managing all those tools used in the Lego approach—to optimize your web experience, collect and analyze web data, connect to other channels, manage tags, conduct ETLs in the data transports from the segmentation tool out to the channels. If you extend this model out, you can easily see where the system starts to buckle, and the outcomes are not what they should be.

The irony, however, is that each individual in the process might think they're doing a great job. The person sending those A/B test, for instance, feels successful as they consider the results of the tests. But without truly connected data, it is virtually impossible to tell if those were the right A/B tests or those A/B tests are advancing the overall marketing cause and improving return on investment.

Each brick in the Lego approach is working as independently based on its scope of visibility, potentially creating discord in the customer experience. Plus, many of the success metrics in that approach are operationally focused versus focused on marketing outcomes or ROI. This situation conceals the strategic shortcomings hidden behind the technical limitations of the environment.

Each brick in the Lego approach is working as independently based on its scope of visibility, potentially creating discord in the customer experience.

redpoint™

Transform with utmost confidence

The Lego approach also creates a significant data security problem, starting all the way back at the data lake and through to the front of the website. With all the sensitivity around data security and privacy, more and more companies are requiring full-time encryption of their data, even while it's being used. (This is called homomorphic encryption.)

Additionally, PII vaults are coming into vogue. This separates PII from operational data, so it resides in a secure location while the other data is used for marketing operations. This approach can be complex and introduce latency issues.

Companies that send out their data to improve identity resolution face another data privacy issue: When you match your data with many software-as-a-providers, you may get new keys every time and there is no way of assigning permanent keys. This makes tracking customers very difficult, so learning about them over time to improve the relevancy of communications and offers because nearly impossible. It also makes it

harder to opt out customers and "forget" them, if they request it.

All of this is causing a lot of nervousness around the lack of data governance around data lakes and code-based approaches. Even retailers are starting to move their personalization from the front of their website to server-side personalization and control. Fortunately, an enterprise-class CDP such as the rgOne platform can handle front- and backend data for personalization.

5 Must-Have Business Transformations

Machine learning, analytics automations, and agile marketing approaches are essential to today's in-the-moment customer interactions. Here are five transformations, recommended by McKinsey & Company, that use these approaches to improve business and marketing performance.

01

Holistic data:

Unify data, as well as data architecture and governance, to enable next-generation advanced analytics.

02

Multichannel customer experience:

Redesign high-impact customer journeys from end to end, including internal and external touchpoints.

03

Revenue management:

Use strategies such as automated cross- and upselling and dynamic pricing and promotions to optimize revenue.

04

Future-ready operations:

Create efficiencies and improve effectiveness using AI and advanced analytics throughout the business—especially to support agile marketing and lean-process approaches.

05

Analytics talent:

Hire and train staff on the capabilities and skills needed to get the most from AI and advanced analytics.

Where we need to go

The Lego approach is fine for small businesses sending out occasional email blasts and doing light personalization on their website. It enables small companies to start in limited ways and grow their way to a larger capability. However, enterprise organizations cannot operate this way; it's akin to stringing together 20 mobile apps to manage your supply chain. Enterprise organizations need enterprise-class software for marketing.

Enterprise organizations need enterprise-class software for marketing.

So, what does enterprise-grade marketing software look like? First, it embraces the data challenge and provides the appropriate level of precision processing automation, cleansing, matching, de-duplication, and data mastering needed to know everything that is knowable about the customer. It enables companies to build a complete Golden Record that links together all of the proxy identities for each and every possible customer—even unknown customers—and provides a robust long-tail of transactional information that includes everything from granular behavior to KPIs to transformations summaries; everything that is needed to know the customer and properly treat and message the customer with exceptional relevance. In addition, all of that data needs to be ingested and processed, and those Golden Records updated, in milliseconds.

The result is a marketing data store that has a complete contact graph and an extensive data story that is valid and current up to the millisecond. Nothing short of this level of data perfection is suitable for a large enterprise. The piecemeal Lego approach is unmanageable at the scale and complexity of large organizations. It would take an army of coders to provide data governance,

do all of the data processing—everything from ingesting the data to cleaning and formatting it, to structuring and redefining it, to standardizing and deduplicating it— to provide what a robust enterprise-grade data processing platform such as Redpoint's rgOne no-code data management platform.

What about all those data scientists that build models using code? A strong data science team is essential to every enter-prise organization and can focus on the high-value models that are crucial to the health and strategic differentiation of a brand. In the meantime, enterprise-class tools such as the rgOne platform's Auto-mated Machine Learning (AML) can gen-erate models that help reduce the friction of the marketing communications process; in the case of rgOne, based on marketers' responses to a set of templated questions.

The automated models allow you, for example, to determine the best time of day to send an email, or the preferred channel of a particular customer, or the preferred next best product for a specific customer. Certainly, you don't need to dedicate an army of highly skilled data scientists to this type of analysis.

Most important, an enterprise-class customer data management platform enables a single-brain approach (i.e., one point of operational control for all channels and messages).

By combining the data management platform with an omni-channel messaging application a single point of operational control and a single point of data control are created that acts as the single brain in charge of all communications. It eliminates fragmented communication that goes out based on local brains (distributed in the channels) that only have a local scope of data. Instead, you can control the utilization and orchestration of all channels, easily selecting the optimal messages, timing, and frequency from that single point of operational control. It enables marketers to strategically engineer a holistic engage-ment journey for individual customers at scale.

This is what a large enterprise needs to properly manage the brand experience— particularly in high-engagement brands such as retail, financial services, telco, healthcare, and travel and transportation. And this is exactly what the rgOne platform

provides: a single point of operational control across all of your channels, delivering real-time updated personalization to virtually every channel that a customer might interact with to communicate with your brand.

4 Types of Customer Data Platforms

Research firm Gartner classifies CDP vendors into one of four groups, based on their capabilities. **rgOne is classified as a smart hub, one of the few that are marketer-driven and empower marketers with rich, robust, accurate data.**

IT DRIVEN ← →   **MARKETER DRIVEN** ← →

1   **2**   **3**   **4**

CDP engines and toolkits	Data middleware	Marketing cloud CDPs	Smart hub
Not a full CDP, these vendors, often open-source, feature resources designed for IT-led teams that are building new applications to complement an existing CDP.	These tools use a marketer-friendly interface to provide oversight of event data streams. But marketers have limited control over delivering segments to downstream marketing touchpoints, and other applications have to handle analytics and decisioning.	Some enterprise software companies that provide marketing cloud platforms also offer a CDP application designed to enhance the marketing cloud functionality.	These CDP vendors emphasize enterprise-level marketing capabilities, including end-to-end campaign orchestration, real-time personalization, advanced customer data analytics, and a single point of control for all event-triggered and planned campaigns and journeys. These solutions enable marketing teams to focus on strategy because the CDP fully automates everything from segmentation to campaign execution.

redpoint™

Many large enterprises are still trying to take the Legos and snap them into enterprise-grade service. The results include frustration, distrust of vendors, and apprehension about what they actually should be doing to achieve their customer engagement and customer profitability goals. It also challenges brand leaders to devise a strategy that works around the technical limitations of a cobbled-together system that relies on a data lake.

But many other enterprise-class organizations are adopting a robust enterprise-grade data processing platform, such as Redpoint's rgOne no-code data management platform. These firms are leading the way with an exceptional level of enterprise communications and customer engagement. The predictable data security and data perfection of the rgOne platform enables large enterprises to engineer their customer experiences down to millisecond of the most impactful moment of truth for each and every customer, and at scale. That level of relevance is what customers expect today, and what the rgOne platform delivers for enterprise-class organizations.
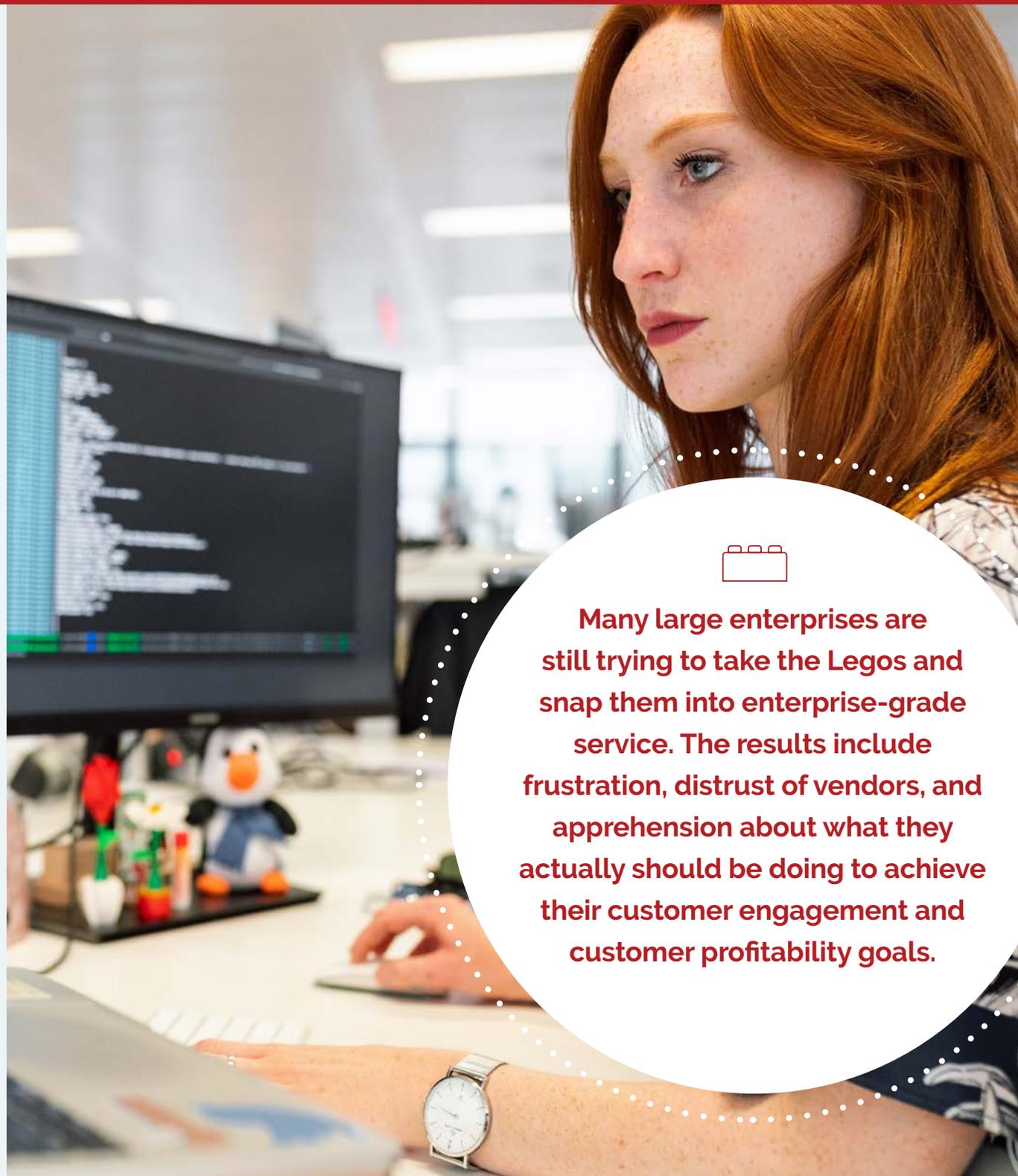
Many large enterprises are still trying to take the Legos and snap them into enterprise-grade service. The results include frustration, distrust of vendors, and apprehension about what they actually should be doing to achieve their customer engagement and customer profitability goals.

redpoint™

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**.

Connect with us

redpoint™

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

EB-DATAUS0521-01