StreamSets
A SOFTWARE AG COMPANY

RESEARCH REPORT

Marketing's Hidden Problem: Data Integration Friction

Introduction

The Data Dependency Pressure Cooker

Enterprise marketing teams are under more pressure than ever to demonstrate value. According to a study by LinkedIn, 77% of CMOs feel under pressure to prove their campaigns are providing a return on investment (ROI). This is increasing the burden on marketing teams to create successful data-driven campaigns that drive conversions, generate leads and deliver ROI.

At the very core of revenue-driving campaigns is data. Marketing teams are dependent on data to meet their operational goals. Without it, they can't effectively segment their audience, identify shifting consumer behavior, understand what makes a successful campaign, or get clarity on their customer pain points. Data leaders and practitioners know that to drive ROI, they must get quality data at the right time, without delays.

But this is easier said than done. Today's data ecosystem is complex and dynamic – it's also constantly evolving as data architectures become increasingly fluid. Building pipelines that connect marketing data from source to destination requires rules to integrate, transform, and process data across multiple environments. Moreover, the data supply chain is not fixed. It shifts from cloud applications and services all the way to on-premises mainframe and legacy systems. All of this makes the job of building resilient data pipelines considerably harder.

It's also further complicated by the fact that managing data and building pipelines are typically IT's responsibility. Marketing teams can struggle to get the data they need to inform their campaigns and create value. They want to do more and achieve better results, but their destiny is out of their hands. As a result, marketing teams sometimes feel forced to act independently.

To lift the lid on the hidden problem of data integration friction and find out what it means for today's marketing data leaders and practitioners, we surveyed data decision-makers and practitioners in marketing. We surveyed large enterprises in the US, UK, Germany, France, Spain, Italy, and Australia to understand the challenges around data integration. In this report, we explore the results and shine a light on the challenges marketing teams face around data access and usability.

What's inside:

- Demand for data is outstripping supply
- Data chaos is holding marketing back
- Beyond friction: cracks in the pipelines
- The true cost of data integration friction
- Governance in the data wild west
- Enterprise-wide collaboration needed to tackle data friction
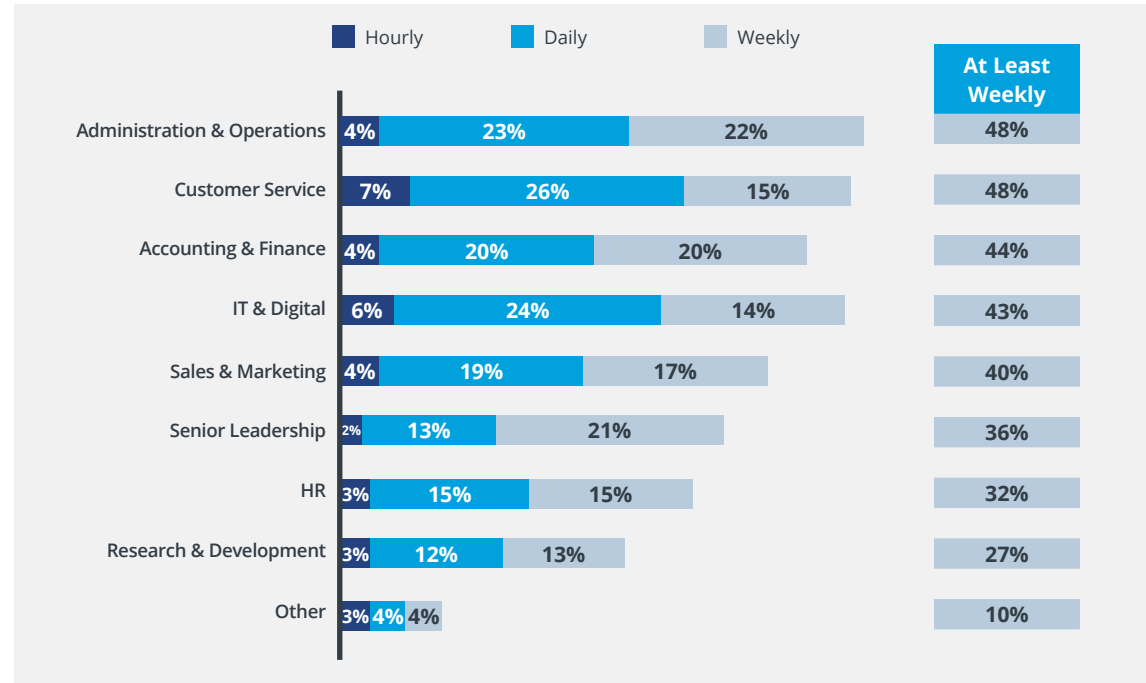- Methodology and demographics

Marketing's Demand for Data is Outstripping Supply

Access to data is critical to every aspect of an organization's digital and strategic objectives. This is especially true in marketing. In a bid to create great campaigns, marketing teams are on the hunt for data. The good news is that they have so much of it at their fingertips, with sources such as CRM systems, website analytic tools, mobile app usage data, social media, and ecommerce platforms. Whether driving new leads, increasing conversions, or improving web traffic, marketing teams require real-time and accurate data to make the best possible decisions and deliver ROI.

Given the pressures they face to improve a business's bottom line, it's perhaps unsurprising to see that sales and marketing teams are one of the biggest data consumers, with 40% of all data leaders and practitioners getting weekly requests from marketing teams.

But this volume of requests from across the business has created a supply and demand problem – marketing teams must compete with other lines of business for attention. Digital transformation is driving demand, creating an insatiable desire for data across all departments. Over half of marketing data professionals (58%) say the acceleration of digital transformation priorities has created major data supply chain challenges.

Figure 1.
How often Line of Business teams request data to support their operations

	Hourly	Daily	Weekly	At Least Weekly
Administration & Operations	4%	23%	22%	48%
Customer Service	7%	26%	15%	48%
Accounting & Finance	4%	20%	20%	44%
IT & Digital	6%	24%	14%	43%
Sales & Marketing	4%	19%	17%	40%
Senior Leadership	2%	13%	21%	36%
HR	3%	15%	15%	32%
Research & Development	3%	12%	13%	27%
Other	3%	4%	4%	10%

The total for each bar may add up to one over or under the weekly column on the right (assuming due to rounding from the original data)

The complexity of enterprises' ecosystems compounds the problem of meeting demand for data. Data engineers must take many steps to connect, transform and process data to build pipelines that meet the individual needs of different departments. But when data is siloed in multiple systems with inconsistent formats, creating bespoke data pipelines at scale is a huge challenge. In turn,

the majority (69%) of data professionals in marketing say this complexity and friction have a crippling impact on digital transformation.
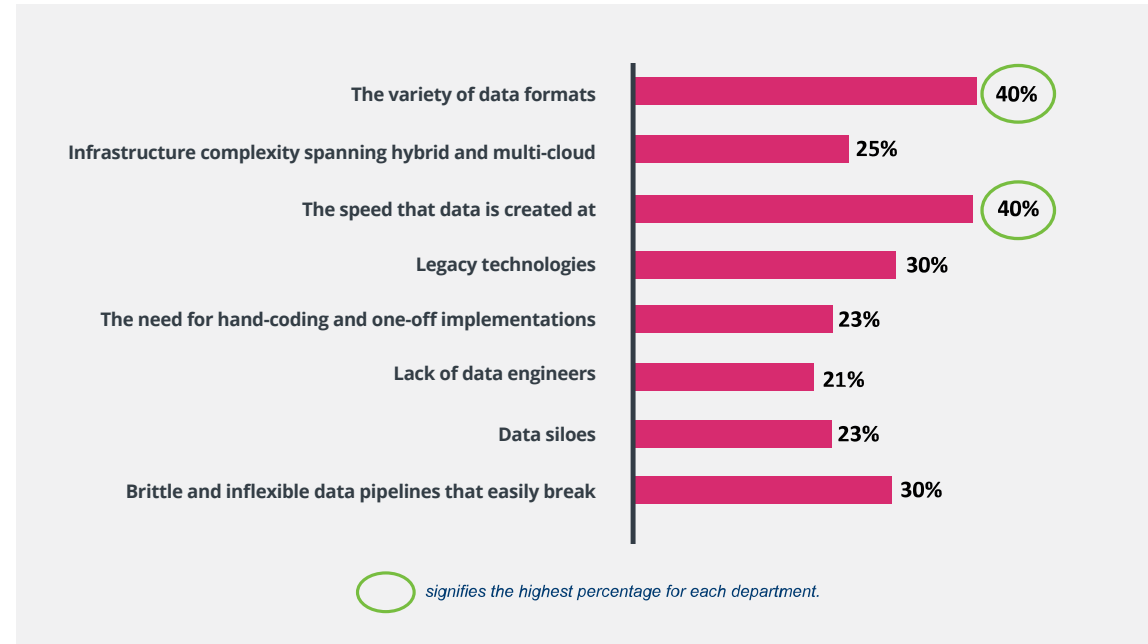
This friction creates a disconnect between the expectations of those in marketing with what can be delivered. Marketing end users expect data like audience segmentation insights on demand and naturally become frustrated by any delay. But non-experts understandably have little appreciation of the scale of the data integration challenge. Over seven-in-ten (71%) data leaders and practitioners in marketing are frustrated that non-data experts think you can click a button and data "magically appears."

The bottom line is that it is difficult to quickly fulfill requests against a backdrop of complexity with scarce resources. Marketing data professionals are under pressure to deliver data and support marketing objectives. Meanwhile, the skilled employees needed to build data pipelines are in short supply. Data will continue to grow in volume, complexity, and urgency. If the data integration friction problem is left unresolved, the inability to empower marketers with the data they desire will impact the effectiveness of future data-driven campaigns.

Data Chaos is Putting the Brakes on Marketing Success

To mitigate data integration friction, marketing data leaders and practitioners have sought "enabling" technologies, which often means moving to the cloud and adopting SaaS tools. But, keeping up with the

Figure 2.
The causes of data friction within respondents' organizations

Cause	Percentage
The variety of data formats	40%
Infrastructure complexity spanning hybrid and multi-cloud	25%
The speed that data is created at	40%
Legacy technologies	30%
The need for hand-coding and one-off implementations	23%
Lack of data engineers	21%
Data siloes	23%
Brittle and inflexible data pipelines that easily break	30%

◯ *signifies the highest percentage for each department.*

constant change introduced by these technologies is hard. It can add to data integration difficulties. As the variety of data marketing teams draw on – and the tools they use to create, deliver, and analyze campaigns – proliferates, it creates a patchwork of systems where data is siloed. Whether legacy systems, point solutions, custom-built tools, or solutions from a cloud service provider, the result is a fragmented and chaotic data environment. This turns what should be a simple data pipeline-building task into a complex job requiring expensive expert skills.

The research found that 73% of marketing data

professionals say data integration friction is preventing them from delivering data at the speed required. Marketing data professionals experience much greater challenges delivering data than peers in other departments, with 30% saying they face major challenges. This compares to 10% in finance teams and 11% against overall data professionals. Those in marketing are also more likely to say data friction is a "chronic problem" in their organization (52%), compared to 43% overall.

There are several factors contributing to this friction. The most cited issue by respondents was the variety of data formats (40%) and the speed at which data is created (40%). Marketing data professionals were also much more likely to cite brittle and inflexible data pipelines as being a cause of friction (30%) compared to 15% overall.

Almost three-quarters (73%) of marketing data professionals also say data in legacy systems, such as mainframes or on-premises databases, are hard to access for cloud analytics, so they often "don't bother" to include it when creating data pipelines. This is a considerable risk. These legacy systems can contain decades of valuable customer insights such as purchasing habits. Many analytics tools today cannot wholly extract data from complex multi and hybrid cloud environments, let alone data trapped in legacy systems.

Marketing teams cannot afford to ignore data stored in legacy systems. It could hold the clue to understanding what customers value and delivering

The Human Impact of Data Integration Friction: Stop Apologizing for Your Data

Trusting the data you use is essential. Analyzing marketing campaigns, understanding consumer behavior, trying to identify trends to exploit...imagine undertaking this work **with complete confidence** in the data you are using.

No one wants to have to apologize for their data. Or justify why the recent website usage numbers aren't included. Or caveat their datasets with explanations of why they only have data on last quarter's marketing campaigns.

This kind of hemming and hawing makes recommendations and resulting campaigns less powerful and can severely damage a marketing team's reputation with the C-Suite.

Yesterday's data is not the same as today's. Marketing teams must be powered by resilient data pipelines that automatically ingest the most up-to-date information and serve it on demand, wherever it is needed. This gives marketers **complete confidence** in the trustworthiness and accuracy of data, so they can stop apologizing for it.

campaigns that hit the right mark. This data is a marketing team's "secret sauce," giving them the edge over competitors. Getting to this data is hard, but it could provide the key to delivering marketing ROI.

Without knowing whether all data sources have been included, marketing teams cannot fully trust their data, else they risk delivering ineffective campaigns based on incomplete insights. Data must come without caveats. So however chaotic the data

ecosystem, marketing data professionals need to be able to run dynamic data pipelines to unlock insights that drive value.

Beyond Friction:
Cracks in the Pipelines

The chaos outlined in previous sections has made building resilient data pipelines hugely difficult. For many marketing data professionals, establishing a pipeline is labor intensive and requires expert support to hand-code one-off solutions that can't be re-used. These manually built pipelines also aren't insulated from unexpected shifts in the environment, resulting in pipelines vulnerable to breakages.

The research found that 60% of marketing data professionals admit their pipelines are too brittle and crack at the first bump in the road. This is significantly higher than the overall (39%), likely because of marketing's propensity to act independently and try new things. Under intense pressure, they are working fast, and pipeline best practices can get lost.

Another reason marketing pipelines break easily is the velocity and volume of data. With a constant stream of data on everything from clicks and payments to sales and conversions, it can be hard to keep on top of all the sources of data available to marketing and the many pipelines in play delivering it to those who need it. This volume and velocity of data can make it challenging to fix breakages, while 53% of marketing data professionals struggle to fix data pipelines in motion.

Almost two-thirds (62%) also say their ability to tackle broken data pipelines lags behind other areas of data engineering. It's perhaps unsurprising that more than nine in ten (92%) have experienced data pipeline

Figure 3.
**How often respondents say their data pipelines break**
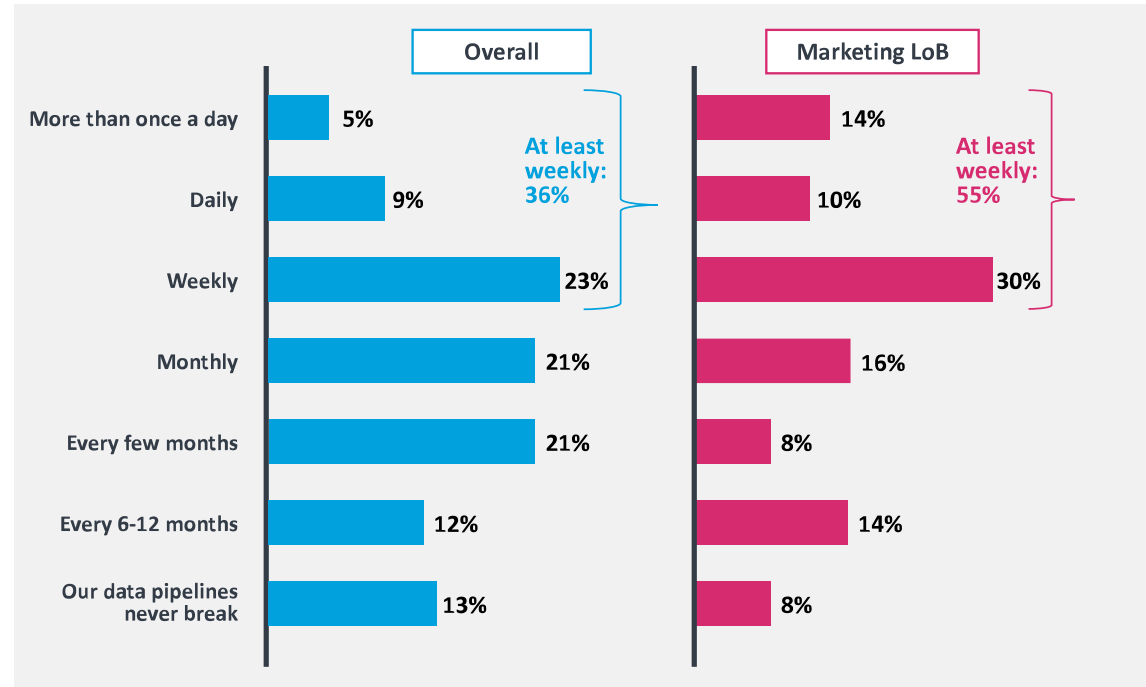
Individual bars may add up to 1% over or under the "At least weekly:" amount due to rounding.

Pipelines break when they are not resilient to changes in the environment. The most cited reasons for breakage by marketing data professionals include bugs and errors introduced during a change (42%), pipeline owners moving on without making a pipeline available to the rest of the team (35%), and infrastructure changes such as moving to a new cloud (34%).

breaks at least once a year, with more than half (55%) saying their pipelines break every week. And worryingly, 24% say they break at least once a day. This is much higher than the average, where 36% of all data leaders and practitioners say their pipelines break every week, and 14% at least once a day.

Without laying these data foundations, every change to how data is stored and used increases the risk of disrupting the data flow to marketing teams. This data drift - the unexpected changes to data structure, semantics, and infrastructure - can break processes and corrupt data, disadvantaging any organization not getting the basics right.

Figure 4.
The most common reasons for data pipelines breaking

signifies the highest percentage for each department.
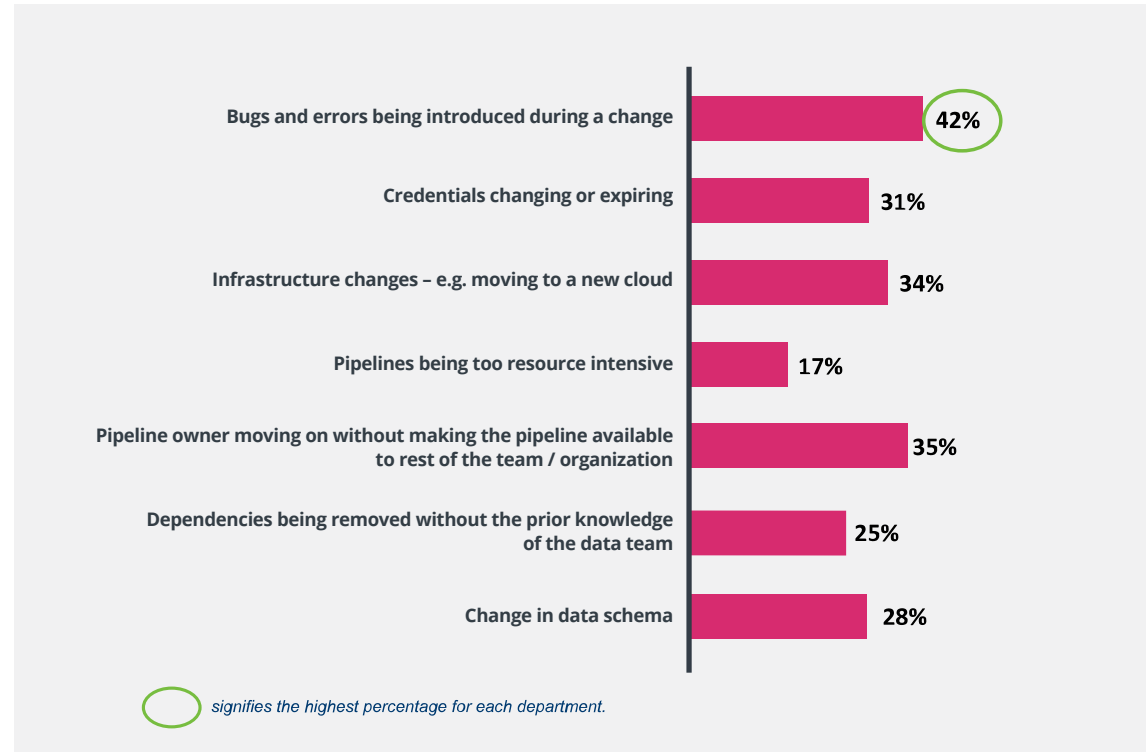
It's hardly surprising to see infrastructure changes high on this list. The adoption of cloud platforms and SaaS applications such as Adobe Marketing Cloud, Salesforce Marketing Cloud, Oracle Marketing Cloud, Marketo, HubSpot, and others can cause problems if marketing teams don't have a clear data strategy for these complex environments. Moving data between these systems requires extensive efforts to orchestrate systems and build flexible data pipelines.

The True Cost of Data Integration Friction

The inability to consistently build resilient data pipelines has major ramifications for marketing data leaders and practitioners. Given the volume of breakages they experience, time spent firefighting swiftly adds up.

Overall, when we look at data leaders and professionals across all organizations, the average data engineer spends 31% of their time troubleshooting and recoding broken data pipelines. When you consider that businesses, on average, spend $6.13 million annually on data experts, repairing data pipelines equates to $1.9 million of their time per year.

$6.13M Average annual spend on data experts

$1.9M Average annual cost of fixing broken data pipelines

This is hardly the best use of data talent that is both scarce and expensive. Organizations must find ways to automate the creation of resilient pipelines and eliminate the need for manual and one-off solutions. Marketing data professionals getting bogged down in this work is a lost opportunity. Instead, they should be free to help their marketing peers identify the data they need to craft campaigns and content that hits

Figure 5.
The percentage of data engineers' time spent troubleshooting and recoding broken data pipelines
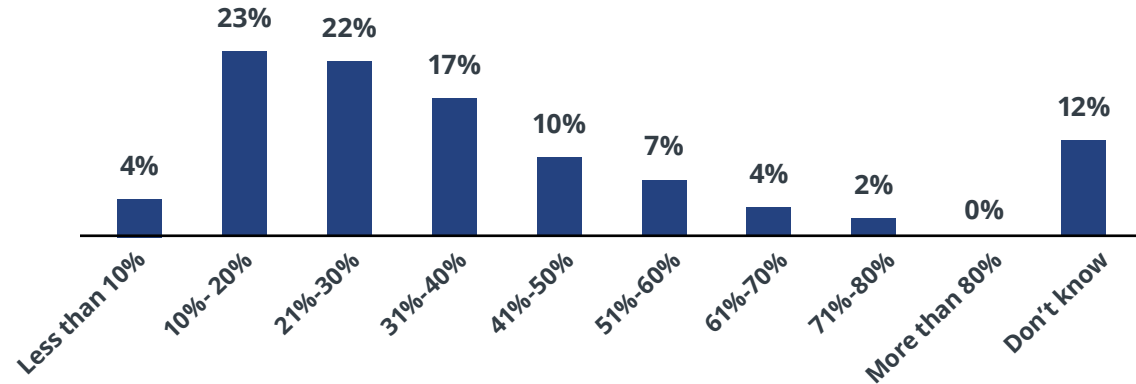
on the challenges their customers face. Only with this can marketing prove ROI.

Unfortunately, data integration friction is hitting marketing's ability to increase revenue. This pressure on the bottom line is something any business wants to avoid.When you consider that businesses, on average, spend $6.13 million annually on data experts, repairing data pipelines equates to $1.9 million of their time per year.

Governance in the Data Wild West

Marketing teams are keen to try new things and adopt new tools to improve their ability to innovate and create impactful campaigns. Even when they

LIFE ON THE MARKETING DATA FRONTIER

56% say that modern infrastructures and data decentralization have created **a data "wild west"**

68% say data fragmentation makes it hard to **understand, govern, and manage data**

81% want **consistent security measures** to protect data as it flows between on-premises and cloud sources

can't get the data they want, they are forward-thinking. While this will benefit the business further down the line, the inclination to act independently also comes with cost – governance challenges and data risks.

More than half (56%) of marketing data leaders and practitioners say modern infrastructures that span on-premises and multiple cloud environments, combined with data decentralization between line of business teams, has created a data "wild west." Further, 68% say this fragmentation in the data supply chain has made it harder to understand, govern, and manage data in their organization.

Connecting the many sources of data available to marketing – be it social media metrics, point of sale data, or ecommerce platforms – has created a complex challenge. For data to be usable, teams need to ensure uniformity between definitions, formatting, and metadata. The need for security and robust governance is also very high.

Regulations such as the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act have put extra pressure on securing data that marketers have access to, such as Personally Identifiable Information. As such, marketing data professionals need a better way to ensure governance over the increasing number of analytics

Five Critical Pillars of Data Governance

Good data governance requires a well-defined strategy. Here are five fundamentals to consider when developing your data governance strategy for marketing.

- **Identify your data:** To design an effective strategy, you need to know your entire data landscape inside out, including types, structures, movements, locations, and points of data transformation.

- **Establish a governance body:** The data governance body is a central control point around which all teams and departments can agree on consistent policies that align with business goals.

- **Ensure "privacy by design":** A privacy-first approach is central to good data governance. It involves collecting only necessary data, masking personally identifiable information (PII), and using data only for intended purposes.

- **Metadata management:** Properly managing metadata makes it easier to track data changes, control data access, and understand relationships between data to fulfill governance requirements.

- **Data quality management:** An effective data governance strategy will establish consistent criteria and scoring to ensure high-quality and reliable data for use in analytics and AI/ML applications.

Once you've designed your strategy following these pillars, you are ready to implement it. Check out **this blog** to learn how.

tools, marketing clouds, and applications that get integrated into the data ecosystem.

The research agrees. 81% of data professionals supporting marketing say they want consistent security measures to protect data as it flows between on-premises and cloud sources. Without consistency, visibility and control are lost, greatly increasing the

risk of data breaches and falling foul of regulations such as GDPR.

Enterprise-wide Collaboration Needed to Tackle Data Friction

Data is the lifeblood of the modern marketing department. It provides the insights needed to create

successful, data-driven campaigns that generate leads, deliver conversions, and prove marketing's ROI.

But as this report has demonstrated, marketing is being held back by data integration friction. And it's suffering much worse than its peers. Given the volume of data available and the variety of cloud, SaaS, and analytics tools at their fingertips, data integration friction can represent a huge disruption. It leaves marketing end users working with outdated information — and flawed data can impact the customer experience and more — while marketing data leaders and practitioners deal with a growing mountain of repair work. Spending time on these repairs prevents them from keeping up with "need-it-now" demands from their marketing peers. Marketing data professionals are aware of the scale of the problem, with 66% believing that smarter data pipelines would enable them to deliver data to the business at pace.

For marketing, there is a pressing need to collaborate with IT and the rest of the organization to reduce data friction. There must be an enterprise-wide effort to unleash the power of data instead of keeping it siloed and sitting with a limited group of experts. It's time to enable more line of business teams to collect last mile data and analyze it themselves. And respondents agree; 94% of marketing data professionals would prefer line of business teams be empowered to do this independently, while 88% want a self-service data model.

However, this world of self-service data collection and analysis requires comprehensive management. End-to-end visibility is critical because every environment has unique deployment and governance challenges. So too is the ability to leverage data from different environments securely and vice versa. 81% of respondents agree, saying a single platform that can handle the complexity of data spanning across cloud and on-premises worlds would be a huge benefit.

To achieve this, data leaders and practitioners across the enterprise must work together to create a centralized management console that acts as a data "mission control." Those who can establish a centralized data console that embeds good governance throughout the organization will ensure line of business teams, including those in marketing, can extract maximum value from data.

This is how StreamSets helps organizations create order from chaos. Our single, fully managed, end-to-end platform becomes an organization's mission control. With StreamSets, businesses and LOB teams have centralized guardrails that allow users in marketing and beyond to explore the art of the possible with the confidence that data is compliant, secure, and up-to-date.

StreamSets helps organizations modernize data integration for continuous data under constant change. It does this by insulating data pipelines from unexpected shifts, so marketing teams can operate effectively in the face of change. By eliminating data integration friction, data leaders can power self-service modern analytics for business teams,

accelerate digital transformation, and unleash the power of data across the enterprise.

Methodology and Demographics

The survey was commissioned by StreamSets; it was conducted among decision makers in marketing for data tools and practitioners who use data tools in the UK, US, Germany, France, Spain, Italy, and Australia. The interviews were conducted online by Sapio Research using an email invitation and an online survey.

About StreamSets

StreamSets, a Software AG company, eliminates data integration friction in complex hybrid and multi-cloud environments to keep pace with need-it-now business data demands. Our platform lets data teams unlock data—without ceding control—to enable a data-driven enterprise. Resilient and repeatable pipelines deliver analytics-ready data that improve real-time decision-making and reduce the costs and risks associated with data flow across an organization. That's why the largest companies in the world trust StreamSets to power millions of data pipelines for modern analytics, data science, smart applications, and hybrid integration.

streamsets.com