The Dirty Secrets of Dirty Data

Data in the New Age of B2B Marketing

CaliberMind

What's On Tap

PART ONE

Marketing's Rising "Data Debt"

An influx of ops teams and data scientists has emerged within the B2B enterprise marketing scene over the last decade. We needed dedicated experts to make sense of the vast amounts of information culled around our customers: Who they are, how they engage, and what they want.

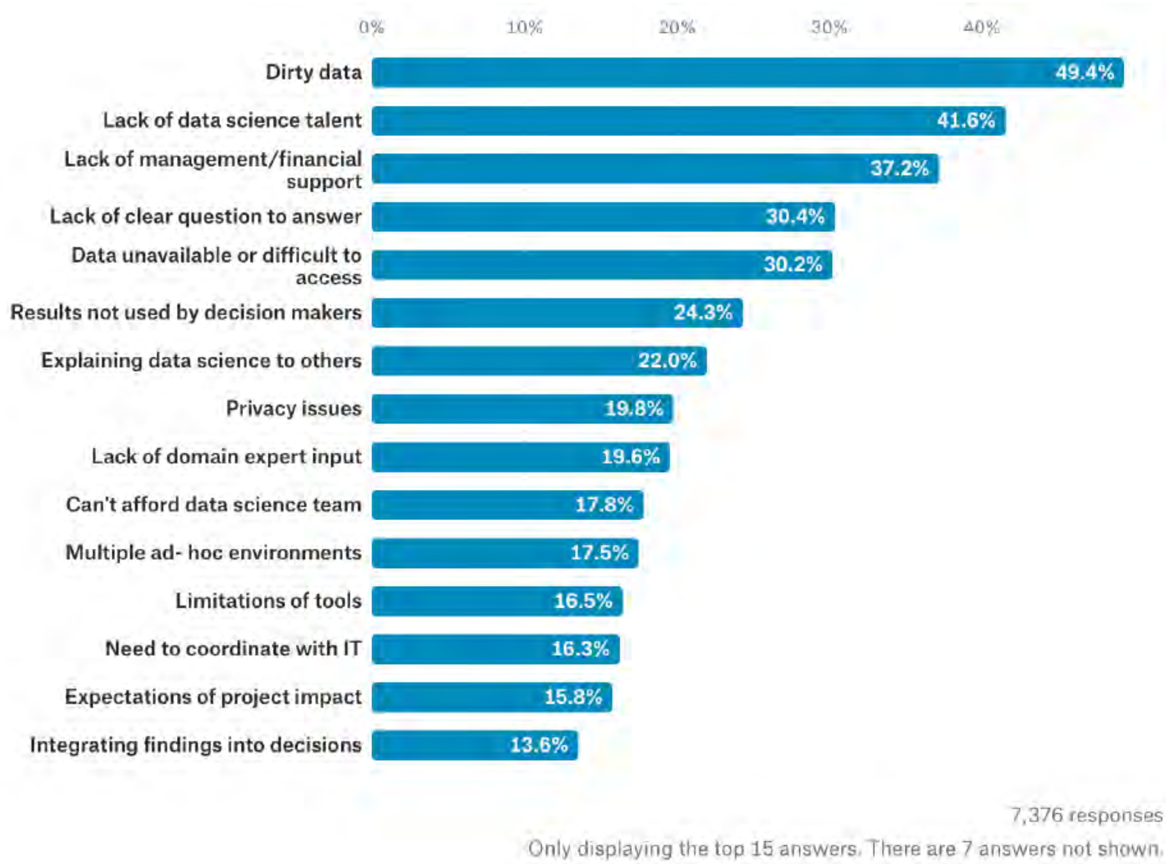
Those well-intentioned and highly efficient minds, combined with tracking technology and faster processing, forged ingenious methods for managing data as it flowed into our systems and platforms.

And flow it does, from dozens of apps, legacy platforms, newly-acquired CRMs, and Marketing Automation Platforms (MAPs). But anyone who's observed a river in the fury of a storm knows that a deluge of rain begets churning waters, muddied by the sudden appearance of thousands of rivulets, winding down valleys and hillsides, bringing with it undecipherable debris and — you guessed it — dirt.

So it goes with every byte we collect in the name of knowing everything we can, and with the desire to attribute X action to Y outcome.

> **Marketing's dirtiest secret?**
> **Our data is full of dirt, and it eats into everything we do.**

We aren't just talking about a little harmless dust. Dirty data is, by far, the most common challenge we face. A 2017 Kaggle survey on the State of Data Science and Machine Learning reveals dirty data as the number one barrier faced among data scientists.



The real tragedy is that before we take a few steps to sufficiently clean up, we all too readily invest in and implement major upgrades. When we skip this step and leave our data unwashed, it will inevitably grime up even the sleekest, most advanced marketing technology.

This contributes to major “data debt.” According to a recent press release by MountainTop Data, **only 13 percent of marketers are confident that they're able to fully utilize the data they get from their MarTech tools**. So, we’re paying enormous sums of money to acquire data, only to render it unusable or unreliable because our discipline around data management isn’t what it should be.

If you’re serious about getting ahead of the attribution curve, or even just designing reliable data reports, take these steps now to get your data in order first. Then, make it shine.

PART TWO

Dirty Data: Identifying the Dirt

A gamut of data enrichment tools exist to collect and clean your data, and often the very notion of managing them all is too overwhelming to consider. You may be thinking, **"I'm overwhelmed with the amount of dirty data I have. How do I even know where to start?"**

Well, the reality is that it's not a one-and-done action, unfortunately. But, you have to start somewhere, and the good news is that the heavy lifting is at the beginning.

After you've cleaned it up once, if you keep on top of your data management consistently, it'll never be that difficult again.
The most important piece to start with, really, is to know what type of dirty data you have, as not all databases are dirty in the same way.

So, let's take a look at the different types of dirty data, and then we'll help you figure out what to do about it.

1 Duplicates -- Leads, Contacts, Accounts

Potentially the most common type of dirty data is duplicate records.

Think about it - how many times have you gone to download a piece of content, and used your personal email instead of your business email?

Then, you download something else from them later, and use your business email, not remembering what you used before.

Just like that, you're a duplicate.

This happens all the time. Accounts are tricky, too, because the main way they are deduped is by your CRM or MAP being able to identify and match domains. If there are two account records in your database, each with a different domain, your systems think they're two separate accounts.

Say your sales rep creates an account manually, not knowing there's an existing account in the system, and misspells the domain.

Duplicate account.

Say a company gets acquired… how does your CRM know to update the domain? Well… it doesn't.

And those are just a few examples of ways duplicates are created.

2 Orphan Leads and Opportunities

Often times, there will be leads in your CRM that never convert to becoming associated with an account as a contact. This can happen if your systems aren't set up for Lead-to-Account matching (which can require development work if you aren't using a CDP or ABM platform), and furthermore require primary keys for a successful match (website, HQ address, email domain, DUNS Number, etc). Something as obvious and intuitive as a "Lead-to-Account" match can be more elusive and difficult than it seems to do consistently, quickly, and accurately.

Leads not matched to accounts are known as orphaned leads.

Without knowing what account a lead is associated with, not only will your Sales team be underprepared when speaking with them, but your reporting could be wildly inaccurate. We all know that there are multiple buyers and stakeholders when making technology purchase decisions. If a lead is engaged but isn't associated with an account, you could be missing key prospects that are ready to buy. Not to mention, if there are other engaged leads from the same organization that aren't being identified as part of the account, your engagement scoring could be off, causing you to miss potential 'hot' prospects.

Orphan Opportunities are exactly what they sound like - opportunities not associated with an account.

Given what we just shared with you regarding Orphan Leads, we're sure we don't need to tell you why these are detrimental to your reporting and attribution.

3 Missing, Inaccurate, or Abnormal Data

Imagine you have a whole bunch of leads and contacts in your database, but a portion of them have inaccurate phone numbers. Or inaccurate email addresses. Or, almost worse - personal email addresses. How are you expected to effectively market to these people?

Not only does this impact your ability to reach these people, but it could also have a significant impact on your deliverability and spam rates if you are trying to reach them through marketing efforts but they keep bouncing back.

Not to mention, this is another way that leads and contacts become duplicates -- if their email address is misspelled or inaccurate in one record, it can't be deduped with the other.

Double whammy.

Another common challenge is data normalization - both newly acquired as well as existing. If you're purchasing data from a third party data provider, this is especially important, as not all sources format data the same way.

Take job titles, for example. They can be all over the place, particularly if your forms have this as a write-in field (think "Manager" vs "Mgr" or "VP of Marketing" vs "VP of Mktg" vs "Vice President of Marketing" vs "Vice President of Mktg" vs "Vice Pres of Mktg" .... We could go on and on).

PART THREE

Exposed Data: Give it a Bath

Okay, we've focused enough on the dirt. Now let's clean it up. Back to all those tools out there that can help you clean and manage your data - it's important to know each tool's strengths and limitations.

Clearbit, for example, relies mostly on email domain to retrieve corporate contact information and details, whereas DiscoverOrg can often do the same job using a job title plus company name, without needing an exact name or email address.

There are plenty of data enrichment vendors out there you can work with to purchase lead, contact, and account data, which is great and all, but how can you ensure that...

> **This information is being updated effectively?**
>
> **All of your systems are talking to each other, regularly, without needing to run a new report every day?**
>
> **And what about those pesky orphan leads and opportunities that a data enrichment provider can't help with?**

This is where a CDP can really work its magic. A relatively new player in the marketing world, a CDP is unique in its ability to tie these tools together, along with the CRM, MAP, and other layers in the marketing tech stack. Let's look at how.

Revenue Marketing Report Interview Series

The Data Era Evolution: How It's Changing the World of Marketing Ops

Sara McNamara, Senior Marketing Operations Manager at Cloudera

Sara shares her candid perspective on how data is changing the marketing world we live in and how we, as marketers, must evolve if we want to keep up with the pace of the industry.

Hear from Sara

PART FOUR

Clean(er) Data: Put it to Work

We mentioned before how data management is not a one-and-done thing. With the continuous influx of data, comes a need for continuous upkeep, too.

Thankfully, CDP's make the process a lot easier, with the ability to identify gaps and discrepancies, automate, push & pull, read & write, model data, and prepare data for reporting.

Here are some examples of key features of a CDP, unable to be found in traditional marketing systems.

Automated Reporting Tools

Account Health

Determine 'white space' as it relates to accounts (i.e. how many accounts are missing a domain?), and identify the number of duplicates, orphan leads, and orphan opps. Useful for determining the level of effort required to clean up your database, as well as regular monitoring of your database health.

Contact Health

Similar to the Account Health report, however the ‘white space’ here identifies the number of duplicates, as well as the number of contacts that are missing key data from their records (phone number, email, title).

Activation & Workflows

Data Normalization

All acquired and existing data needs to be categorized (industry, titles, employee tiers, etc). This functionality is built in to a CDP out of the box, with standard templates that work for most organizations. And, for organizations that have custom rules, CDPs are flexible enough to work them in.

*Note: one **very important** normalization workflow to have running is for standardizing and fixing websites on the account record, as this is needed in order to dedupe accounts (more on this functionality below).*

Data Enrichment

CaliberMind's CDP has a number of different workflows built specifically around Clearbit, which enable you to access many of their enrichment capabilities. Use these data enrichment workflows to find important data points about partial data you have on both a person-level (title, social handles, phone number, affinities, etc) and a company level (technographics, firmographics).

Let's take a deeper look at enrichment.

Combined Enrichment

Takes a list of email addresses and enriches both the Account and Contact information. The results are placed into a new list that can then be used to send to Salesforce, a Marketing List, etc

Prospector Enrichment

Enables you to take a list of accounts (by domain) and input titles, seniority and business unit information to find contacts. The results are put into a list that can be used to send to Salesforce or your marketing automation platform.

Name to Domain Enrichment

Will take a list of Account Names and enrich their website domains. This is a useful enrichment when looking to take a list of source companies and eventually enrich the Account information and then specific targets at the company by title.

Pro Tip

At CaliberMind, we use data enrichment as an important first step in everything we do. You may have noticed that on our website forms, we only ask for email address. When we receive this one piece of data, we run a Combined Enrichment workflow to enrich the contact data, and simultaneously identify the account.

Within the same flow, once identified, the account is enriched with firmographic data, which we then run against our ICP scoring formula (which also lives in CaliberMind). So, in one real-time workflow, we're not only able to gather data about the contact and account, but we've also been able to determine how likely that account is to convert to business.

More Activations & Workflows...

Data Deduplication

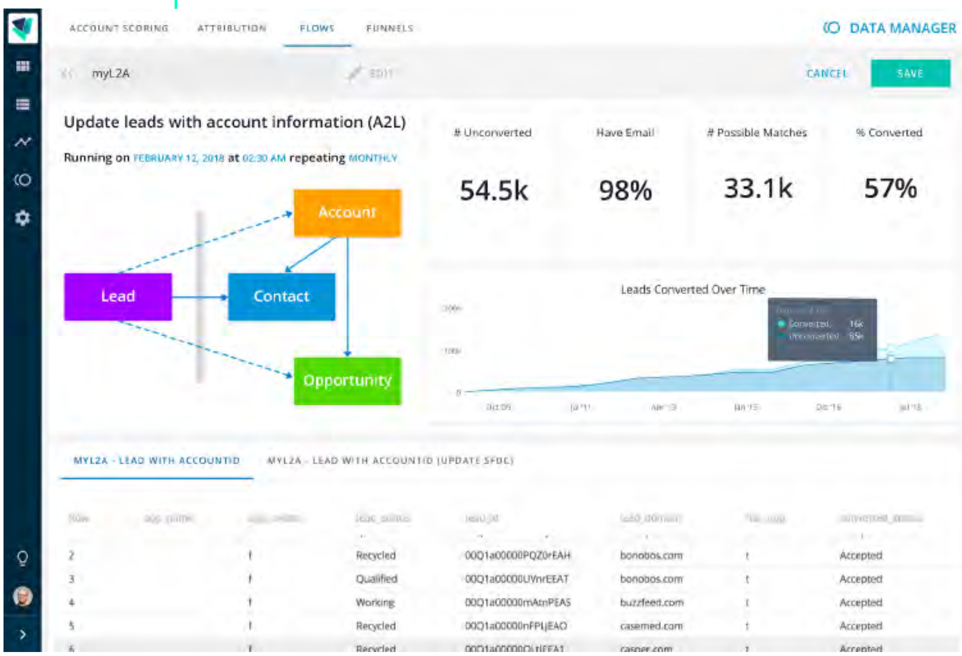
Having a deduplicated database is a critical step in accurate reporting and efficient processes. Thankfully, a CDP makes this very easy. All you have to do is run an automated workflow that looks for matching record information (for contacts, we match on email address, and for accounts, we match on domain - again, this is where Data Normalization is crucial). This workflow will merge any duplicate records it finds in Salesforce.

Lead to Account Matching (L2A)

When systems like Salesforce and Marketo were designed, the B2B world thought in 2 camps of people (Leads and Contacts). These systems were not built to properly manage the concept of "People" and "Companies" out of the box.

But, we also know that account scoring is the new way of identifying companies that are engaged with your offers. A CDP can review a lead’s record, identify their email domain and website, and run through the database of accounts to find one that matches. It automatically associates that lead with the account and converts it to a contact.

Here's an example of the L2A Matching workflow we use at CaliberMind.



Account to Lead Matching (A2L)

Similar to L2A matching, A2L matching identifies a lead's account based off of email domain and website, however it stops before the step of converting it to a contact. This process still associates the lead with the account, but gives your sales reps the option to manually convert to a contact. Since Account Based Marketing can be a big organizational change, A2L is an alternative strategy for companies with very ingrained (Lead-centric) processes.

The Final Set of Activations & Workflows

Data Operationalizing

Read-write into all core systems so the business is aligned and customers receive accurate and relevant communication.

Data Institutionalizing

With a community of users and advanced training, marketers are enabled to take ownership of their data. This also encourages adoption of marketing metrics at the board and CEO-level.

Data Modeling

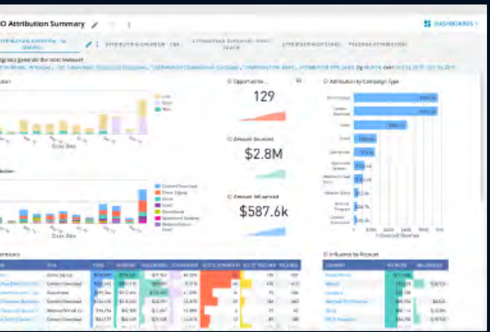
Most analysts are not experts in B2B marketing. CaliberMind has the unique advantage that our analysts are both that, and data scientists. As a result, our product is built using industry standards and machine learning. (ex. Fit Scoring aka Ideal Customer Profile Scoring, Account Engagement/Intent Scoring, etc). Data normalization enables this, by limiting the number of features in your data set.

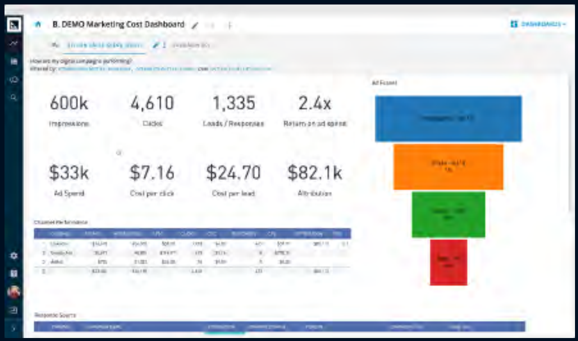
Let’s Talk About Data Visualization.

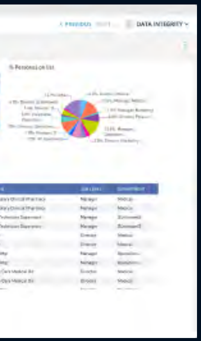
What good is clean data if you don’t have an easy way to report on it to your stakeholders? One of the largest benefits of a CDP is out-of-the-box reporting and visualization capabilities.

CaliberMind has 30+ most commonly requested reports by CMOs and Marketing Operational professionals already built-in from the get-go. Once you’ve gone through the process of ensuring your data is clean, our CDP allows you to easily and accurately report on and visualize key metrics for marketing performance in a digestible format.

Among the categories of reports? **Attribution, Account Based Marketing, Demand Generation, Funnel, and Return on Marketing Investment, and Customer Success.** Read on to learn details on the most widely used reports, and why they’re crucial for success.







CDP Playbook

A card-based, interactive collection of Customer Data Platform plays that includes reports around Demand, ABM, ROI, and Attribution.

THE END

That's a Wrap.

The deeper into data debt we fall, the harder it is to be successful at marketing. Data is growing explosively. With more technology available than ever before, the data we're charged with understanding becomes both our biggest asset, and a liability.

The dirty data behind why attribution fails most marketing teams is something everyone acknowledges, but doesn't or can't take the time to fix.

Marketers, it's only going to get worse. With some initial clean-up effort and regular monitoring, a CDP can enable teams to get beyond the dirt and make better use of the data for truthful attribution measurement, ROI and ROAS reports, and ultimately, higher revenue contribution.

EXPERT B2B MARKETING GUIDE

Dirty Data's Dirty Secret

CaliberMind

CaliberMind is a customer data platform built for B2B enterprise revenue marketers that use data, analytics, and automation to grow revenue faster, together with sales. Learn more at www.calibermind.com.