Talk dirty data TO ME

Data is the new currency of business. But what if your data is incorrect?

By Matt Hooper and Daryl Pitts, Saxony Partners

Although it sounds provocative, dirty data is just data that is incorrect, incomplete, or miscategorized. It is often borne of simple inaccuracies, poor processes, and/or inadequate software solutions.

Dirty data isn't exclusive to any specific type of company. Rather, it is universal — every company has some dirty data, even yours. It is the digital equivalent of a red shirt hidden inside your white laundry; many companies don't spot the issue until everything is pink.

There are ways to identify dirty data, clean it up, and keep it clean. But before we discuss solutions, let's dive deeper into the what, where, and how of dirty data.

Types of dirty data

Where there's data, there's going to be dirty data.

This is particularly true of real estate data, which covers multiple metrics across the entire life cycle of the asset. New data is constantly entering the life cycle from the investors, to the underwriters, to the asset managers, and back to the investors. If that data is compromised, then every subsequent stage within the life of the asset is compromised, as well.

Although any of this data can be dirty, property-level data is likely to be the most compromised.

Lease-oriented data tends to be the dirtiest. This includes everything from inaccurate summations of the lease (reflected in Yardi or some other platform) to misspelled tenant names.

Imagine you own several multifamily housing developments. Now perhaps you have accumulated a few slight variations on the spelling of tenant names, or a misplaced digit in a few phone numbers. Not only does this mean your lease data is incorrect, it means you probably have some duplicate data. When you discover that Mrs. Brice actually spells her name "Bryce" — you likely now have both names in the system. The relationship between inaccurate data and inefficacy is directly proportional. Inefficiency leads to higher employee headcount — someone has to sort through the mess, after all. And higher headcounts translate to lower margins.

[1] Altus Group, *Future-Proofing the CRE Industry: Is Commercial Real Estate's Innovation Curve Moving Fast Enough?*

And then there's dirty accounting data. The team at Saxony has seen this scenario play out more than once: an asset manager, anxious to produce a report or complete a task, bypasses the accounting department and grabs numbers from wherever they can be found. Are those numbers up-to-date or properly contextualized? Who can say? (Well, accounting could say, of course — but they were passed over.)

The result? At best, inaccurate data, and at worst, shadow accounting.

How does dirty data get dirty?

There are innumerable ways to corrupt data, but most common are lack of data governance, a preponderance of siloed processes, and inadequate software solutions.

Data governance is a multi-legged stool, with one of those legs being data integrity. Processes for capturing, cataloguing, and reporting data have to be consistent across your organization. If, for instance, your asset manager defines a certain term differently than the portfolio manager (and neither is aware of the difference), you will have inaccurate data. It's safe to assume that lack of data governance broadly — and lack of data integrity specifically — represents the single greatest contributor to dirty data.

Siloed processes can contribute to that lack of data governance. There's no reason for asset managers and portfolio managers to be out of sync with one another, but it happens often. When companies splinter from realms into fiefdoms, the lords and serfs each feel compelled to devise their own independent processes and procedures.

Say, for instance, that an asset manager wants to execute a tenant exposure report or a top-ten clients report. Pressured by time, she/he decides to create a proprietary, off-line procedure to expedite the process. It works — or at least it works until that asset manager leaves the firm and takes their proprietary knowledge with them.

What motivated the asset manager to go rogue? If the firm was utilizing *inadequate software solutions* to manage asset data, then their reporting was probably more painful and time-consuming than it should have been.

According to a recent survey,[1] roughly three in every four commercial real estate firms are using Microsoft Excel to manage their core investments. For many of these firms, this represents hundreds, perhaps thousands, of spreadsheets spread across multiple locations. Such a scenario leads to two big problems: one, it makes work more complicated and firms less efficient; and two, it increases a firm's organizational and reputational risk.

Excel is popular because it's flexible and widely accessible. But that flexibility is a double-edged sword when it comes to creating a database. Because it lacks rigid structure around data governance, employees have the implicit ability to invent shortcuts that undercut data integrity.

If not Excel, then what? The same survey found that more than 70% of CRE firms surveyed realized the potential of integrated, cloud-based management solutions to manage their critical data. These solutions, like Saxony Partners' Pereview Software, provide a secure, integrated platform for data across the entire life of the asset.

Such software platforms boost data governance by enforcing strict rules around accumulating, managing, and reporting on asset data. And because these specialized software platforms establish an integrated, single-source-of-truth for your firm's data, employees are blocked from engineering their own proprietary processes.

> "There are innumerable ways to corrupt data, but most common are lack of data governance, a preponderance of siloed processes, and inadequate software solutions."

iStock.com/DrAfter123

The effects of dirty data

Doing things wrong increases risk, turns off investors, wastes money, and risks legal repercussions.

If you present a report to your investors stating that their return will be X, when the actual return is Y, there's going to be a reckoning. There will probably be some digging into your actual numbers and your processes. And, depending on what they find, there could be a loss of confidence in their estimation of your firm.

That loss of confidence trickles down from investors to management, and from management to employees. After all, it's the employees who are consistently reporting this bad information.

Of course, issuing erroneous reports increases more than just reputational risk. Dirty data might give the government a reason to poke around into your numbers and processes, too.

Fixing faulty reporting and dirty data requires time and people. Like crime scene investigators, employees have to retrace processes, track down real numbers, and retroactively apply fixes to reconcile accounts. Headcount expands, money is wasted, and morale sags.

Discovering dirty data

Of course, the forensic investigation employees' undertaking of an effort to find and repair damage done by dirty data is usually step four in the dirty data remediation process. Step one: Data gets dirty. Step three: Companies discover dirty data. Step four: Companies work backward to correct the problem. Step two is when someone starts running quarterly and monthly reports using this dirty data.

Step two happens frequently when the onboarding process begins, as is customary, with a business process review. During that review, many of the company's regular reporting is reproduced and compared with previous reporting. Inconsistencies between these reports have nowhere to hide.

> "Cleaning data and keeping data clean are two different things."

Cleaning dirty data — and keeping it clean

When reporting uncovers dirty data, consultants can help you devise processes for cleaning data (and keeping it clean). But it's up to the company to execute that forensic investigation and reconcile their own data. Doing so improves workforce efficiency, increases productivity, mitigates risks, and establishes a firm foundation that allows a company to do more with their data.

Cleaning data and keeping data clean are two different things. You can do the latter by establishing strict guidelines around data governance, implementing standardization, improving internal communications, and adopting software solutions serving the entire life of the asset.

Standardization forces companies to clearly outline data governance processes, definitions, categorizations, and accountability into one playbook. This process should be led by the company's chief information or technology officer and include input from everyone who is directly involved in data governance and maintaining data integrity. Once created, this playbook must be distributed to every department within the organization. Accountability measures outlined within the playbook must be followed to ensure that siloes are not being rebuilt.

Companies must also build standardization into their reporting processes. This is very difficult, if not impossible, to achieve within the Excel environment. Pulling data into a software platform that can collect, secure, and manage data across the entire life of the asset radically reduces the amount of time and personnel needed to execute those regular reports. And having all of your data in one holistic system greatly reduces the chances of compromised data.

Conclusion

Your company is affected by dirty data. Dirty data has, to some extent, impacted your efficiency, increased your risk, muddied your reporting, and impacted your bottom line. That's the bad news.

The good news is that you are not alone — there's hardly a real estate company anywhere that's not facing the same situation. And there's more good news — by shoring up data governance, implementing standardization, and evacuating your data from Excel to a life-of-the-asset software solution, you can ensure that your data is clean, now and in the future. ◆

Matt Hooper joined Saxony Partners in July 2018 and leads the company's strategic communications and branding department.

Daryl Pitts joined Pereview Software as Senior Vice President of Sales in February 2018.