From: www.cio.com

Stupid Data Corruption Tricks: Take Our CRM Quiz

– David Taber, CIO

November 02, 2010

In the spirit of David Letterman's occasional feature "Stupid People Tricks," I thought it was time to summarize common errors that can lead to corrupted CRM records, or worse. How much worse? Read on.

Most of the opportunities for corruption come when doing data imports, data cleanup, and other mass update operations. Many of them come from using standard external tools like Excel, others come from interactions with ETL, data cleansing, and mass-update products.

So here's David Taber-man's Top 10 List of Stupid Ways to Corrupt Your CRM Data...

Number 10: Open a CSV file directly into Excel

The geniuses in Redmond made Excel so easy to use, it automagically corrupts phone numbers and other large integers (turning them into scientific notation) when you open CSV files by double-clicking on them.

Number 9: Copy formulas that use relative coordinates

There are several variants of this problem, but they all lead to insidious miscalculations, particularly if your formulas include error-handling. A variant of this comes when you're intending to copy cell contents all the way to the end of a row set, but the copy-down operation stops at an empty cell that you didn't notice.

Number 8: Accidentally use VLOOKUP's fuzzy match

Excel's VLOOKUP function is really useful, but too often people don't read the fine print about its behavior. Nasty corruptions occur when formulas don't include the word "false" as the fourth VLOOKUP parameter.

Number 7: Put values in fields that are supposed to be pointers or references

In most CRM systems, and in most object databases, many of the data items are actually references to another table's entries (e.g., "account name" in the "contact" record). While the system will immediately complain if you put long strings in those pointer fields, it may not if you put in a short string (like "Acme") where the pointer is supposed to go. Once you've saved this record, it can be tough to identify the references that point to nowhere.

Number 6: Miss the data type

In strongly typed systems, it's impossible to confuse a date or a date time with an integer. But too often when working with tools outside of systems, dates can be misrepresented as integers, long integers as short floats, and similar data errors. This can cause some truly amusing messes.

Number 5: Use a deduping tool with "loose" criteria first

In most CRM systems, merging duplicate records involves some complex internal operations. Consequently, merges cannot be undone. If you use a deduping tool and start with the loosest criteria first, the false-positives can easily be 50 percent of the result set...and reversing the merges is a complete nightmare.

Number 4: Sort a spreadsheet, but not include all the columns

Ease of use hell, revisited. Excel's "select all" only select contiguous columns, so if you have empty columns some of your data won't be included in "all." Sort the spreadsheet, and you've corrupted every record there! Yum.

Number 3: Start working on the database without doing a full backup first
This is like telling the electrician to turn the circuit breaker off — but in a rush, this mistake happens. Even if your CRM vendor has continuous backup for free, that doesn't mean data restores will be gratis. There's a local variant for this one: not remembering to do a checkpoint save (to a versioned file name) from whatever tool you're using every hour or so.

Number 2: "What system am I logged into?"
For integrators who have to work across several systems in a day, this one has truly horrific implications. Update 5000 rows, but do it in production rather than the sandbox system. Or worse, make changes in client B's system that were supposed to happen in client A's. This is one of the best ways to measure the ohnosecond.

And the Number 1 way to corrupt your data...

Click "yes" without carefully evaluating the message that says "do you want to remove this from the server?"
Tools really try to help us do the right thing, but sometimes the dialog boxes just aren't specific enough. "Server" might mean your machine in one instance, or the sandbox in another, or the production system in another. This is the error that can actually delete metadata and configurations, not just data.

Scoring
Of course your team and contractors will say they've never done these — they'll get a score of zero. That means they're either hiding the truth or haven't been working under enough pressure!

And no, I've never done any of these...

David Taber is the author of the new Prentice Hall book, "Salesforce.com Secrets of Success" and is the CEO of SalesLogistix, a certified Salesforce.com consultancy focused on business process improvement through use of CRM systems. SalesLogistix clients are in North America, Europe, Israel, and India, and David has over 25 years experience in high tech, including 10 years at the VP level or above.

Follow everything from CIO.com on Twitter @CIOonline, and the CIO.com Facebook page

© 2010 CXO Media Inc.