The St. Valentine's Day Data Massacre

Ensuring data security, integrity & validity during software development

An Original Insight

Original Software

Introduction

On the 82nd anniversary of the legendary US gangster slaying in Lincoln Park, Chicago, it is perhaps appropriate to use the title of the massacre as an analogy to describe two of the most dangerous, but widespread practices currently taking place within QA and testing departments around the globe.

Illegally testing on live data and ignoring database accuracy!

There are those who argue that verifying software quality by using live data is a necessity as it gives more accurate results. However, using customer, employee, or other confidential data straight from production for testing or developing applications violates data privacy laws and regulations and makes that data a soft target for attacks.

Exposing traceable personal production data to the QA team is unnecessary and betrays the confidence of the data provider as well as breaking compliance laws! However, despite the legal requirements and in some cases severe penalties for breaking them, testing on live data still tends to be common practice.

A study carried out by the Ponemon Institute, "The Insecurity of Test Data: The Unseen Crisis" showed 62 percent of companies were using live customer data to test applications and 49 percent shared this data with outsourced testers, with no way of knowing if it was ever compromised.

The simple fact is this: Using data in a test environment, without the necessary security measures in place, is in breach of regulations and compliance laws, such as EU Data Protection Laws, FDA, FSA, HIPAA, PCI, SOX, etc.

Ignoring what happens in the database itself as a result of the test process, is equally perilous.

The second concern refers to how organisations value their data. Errors in the database are often the most costly and the most difficult to detect, causing hours of time spent tracing hidden problems and at the same time they can be the cause of the most severe and wide-ranging damage. There have been numerous cases of invoices, bills, statements, and charges being raised in error to large volumes of customers because of errors in the data. By the time the underlying data error is discovered, it is already very public and a combination of technical recovery and high profile PR recovery are required as emergency measures.

This is exactly what happened to credit card giant Visa. An empty amount field in the database caused 13,000 card holders to receive warnings that their accounts were overdrawn by more than $23 quadrillion.

So focusing testing solely on the User Interface is ignoring the real risk areas and gives a false sense of security.

So, what can be done? Is there a way for QA staff to carry on testing with live data without leaving themselves and their organisation at risk? In short, can this 'data massacre' be avoided?

This 'Original Insight' will show you how, having a clear Test Data Management (TDM) strategy in place, you can ensure the security, integrity and validity of your data and be happy in the knowledge that your entire application quality process is underpinned by accurate and legal test data.

"…two of the most dangerous but widespread practices currently taking place within QA and testing departments around the globe. Illegally testing on live data and ignoring database accuracy!"

DATA TIP:

Data privacy is not just a concern for production systems; it must extend to your test environments too.

A Test Data Massacre

An unacceptable exposure

Industry estimates suggest that 62 percent of IT departments admit to using live data during their application testing process. One can understand why when you consider the fact that it makes the testing more realistic, gives a clearer indication of true quality and allows the application in question to be tested more thoroughly. However, the problem is that not only is this method slow, it exposes sensitive data to less than sensitive employees or contractors, who are not necessarily authorised to view such data.

According to Forrester Research, only 16 percent of enterprises surveyed in 2008[1] indicated that they perform data masking to support their test environments. A survey by Carrie-Ann Skiner[2] showed that "four out of ten office workers have taken sensitive company data from their employers to a new job and almost half would steal data as an "insurance policy" in tough economic times." This trend is worrying and is all the more reason to ensure your test data is secure, safe and legal.

The use of live data for testing may breach the data privacy safeguards of your live system which could result in fraud, malicious damage or even legal action if data confidentiality is lost or compromised.

In addition, there is the cost issue of data loss. The Ponemon Institute have revealed through their research that data breaches cost organisations an average of $140 per record. That may not sound much, but when you 'lose' a million records, that is an expensive mistake.

DATA TIP:

Just because data is stored in a test or development environment does not release your company from the responsibility to comply with privacy regulations.

The heart of the matter

The database is the core of most applications. It contains the information upon which the application logic depends, processes and probably affects. Yet, it is frequently not tested itself and tested only by proxy on the basis that testing the application window into the data is good enough. This is a position which immediately appears weak and vulnerable as the database itself warrants attention in the test process. It requires significant inspection in order to validate its integrity as well as the performance of the functions affecting it. The ability to combine simultaneous testing of the user interface and the underlying effect on the database would be a powerful combination. This way, both the user interface and the impact on the database can be simultaneously verified. The consequences of insufficient attention to data accuracy are often very costly in both financial and reputation terms.

"It looks to me like somebody blank-filled a field, plopped the actual charged amount into the end (hex 1250, decimal 4688, likely amount $46.88), and then interpreted the entire field as a hex number. If so, this is the kind of bug that would have been caught in even the most cursory testing, in which case the 'technical glitch' Visa talks about was not really about the software bug, but in their own shoddy procedures that allowed untested software to go live."[3]

Ignore at your peril

"But hang on, we have been doing it this way for years, and anyway, we have security procedures in place."

Both these issues maybe time-bombs waiting to waiting to explode at great cost perhaps just in terms of effort to fix, but commonly also in real and large fines from legislators. Just because the bomb has not gone off yet, does not mean it is not a risk. A risk CIOs and Compliance officers must address to avoid professional and commercial suicide.

[1] Forrester and TechTarget conducted a joint survey in Nov 2008 on database management.

[2] A Computer World UK article.

[3] A quote extract from: http://www.theregister.co.uk/2009/07/16/visa_programming_error_cracked/

A Viral Flu Epidemic...

The online world continues to evolve with online shopping and online banking becoming the norm. At the same time Web 2.0[1] and social media networking are taking centre stage.

Sharing news or opinions on the internet has become as viral as a flu epidemic. We live in a world of "tweets" and "blogs", wrapped up in a blanket of virtual deception. Individuals, communities and companies who are embracing this Web 2.0 culture, are tapping into a method of communication that allows immediate and direct access to their peers, prospects and customers. This means that data breaches, errors and exposure become public very quickly, and organisations need to react swiftly to bad publicity and diffuse the situation.

"Sharing news or opinions on the internet has become as viral as the flu epidemic."

In these days of peer to peer recommendations, fortunes can be made or broken by the online community. If a company breaks the law and pays a fine, this can in some cases be far less of a punishment than to be judged and exposed online. It's a bit like the ancient Roman games. Imagine the great Coliseum in Italy, where many Gladiators won and lost. As much as Caesar, (the law enforcer), ultimately passes judgment on the loser of a match, it is actually the people (that's our online users), who pass their own judgment and forces Casers hand...thumb down!

Last year alone, saw a number of blockbuster data breaches. Federal Reserve Bank of New York, Heartland Payment Systems, Camden Primary Care Trust are just a handful of the businesses that fell foul of data security regulations.

The Responsibility that goes with Personal Data

Around the world, regulators are getting tough on data security breaches. In the UK, it is mandatory for all organisations that process or hold personal data to comply with The Data Protection Act. In the US, a bill that would set out the requirements for ensuring data security has been approved by the House Energy & Commerce Committee.

Additionally, the Personal Data Privacy and Security Act of 2009 is a bill that calls for enhanced criminal penalties against security breaches. The UK Information Commissioners Office will now be able to issue fines of up to £500,000 for serious data security breaches.[2] This is an update to the Data Protection Act, which originally came into force in 1984.

Since the evolution of the Web and the growing popularity for businesses to trade online, enormous amounts of personal data is being stored and processed, in turn the law is trying to keep up with the change in technologies to help protect online users.

"Computer Weekly: Better testing may have prevented Parcelforce data breach."

With the rise in online banking and eCommerce, customers are putting their faith into brands and the software they are directed to use. Customers can only assume (and hope) that their personal data is kept safe and everything is okay. But it is only when a story is leaked, that attempts to endanger the faith between the customer and vendor, do we truly see what goes on behind the brand. It is then that we question the security of our data and the integrity of the software we have used.

Stories such as these often end up on the Original Software "Software Quality Hall of Shame". A recent hall of shame article was reported on in Computer Weekly: *"Better testing may have prevented Parcelforce data breach."* The data breach at Parcelforce meant that when customers entered their parcel tracking numbers online, they were able to gain access to other customers' delivery details. In this case, Parcelforce had not tested their site properly and had not ensured data security or integrity.

Nightmare stories of database errors also come up from time to time, such as a news story published in the "Telegraph" about a Swiss bank who had brought the Tokyo Stock Exchange to a halt when it was discovered that a nought in the database was in the wrong place. A £22bn order for convertible bonds was raised instead of the intended £2.2bn order.
Both Parcelforce and the Tokyo Stock Exchange were subjected to bad publicity that was spurned on by the online community!

[1] Wikipedia Extract: Examples of Web 2.0 include web-based communities, hosted services, web applications, social-networking sites, video-sharing sites, wikis, blogs, mashups, and folksonomies.

[2] BBC News article published 12th January 2010.

Test Data Management: Avoiding the Data Bullets

IT departments around the globe have a significant corporate responsibility, holding the key to the organisation's data bank. It is a responsibility that needs to be addressed in depth and conscientiously through appropriate testing and measures to ensure data security, test data integrity and database validity.

Test Data Management (TDM) is one of the fundamental components to the success of your test strategy; after all, data drives the entire testing procedure. But there are a number of straightforward actions which can be implemented to provide protection and enable you to dodge the data related bullets.

Firstly, look at the quality of test data. With bad data comes poor testing, results you cannot trust, wasted time, money and effort. Very often production data is a great source but it has issues which need to be addressed if it can be used effectively and safely. Good test data will be; realistic, succinct, accurate, a good cross section, re-usable, and support test cases. Getting this combination right will less storage space, faster testing, better testing and will provide a platform for automation which can provide enormous time savings. Intelligent data extraction dramatically simplifies the process of creating and extracting data subsets from your live database with full referential integrity. With total control to amend data during extraction, and the ability to extract data from remote sources, data maintenance becomes simple and efficient. Data integrity, for any purpose, is assured.

This means developing a strategy that will select a sample of data from production to form a perfect test-orientated subset, perfect for testing and thus enabling maximum efficiency and quality.

Secondly, the resulting data is still production data, even if now residing in a test environment it is production data and needs to be cared for as such. In particular, personal data has not usually been provided for your software testing use and you also have an obligation to ensure it is accurate and not made accessible to more persons than need to use it. This often prevents companies using production data, but the realistic alternative is to change it, to obfuscate it in the test environment. 'Scrambling' is perhaps the most effective method for this, addressing the needs of the tester, protecting the data provider and protecting the organisation from exposure and financial penalty.

There is also a stage in the software development lifecycle that the integrity of the database needs validation. This is often a technical challenge requiring both detailed application and tool knowledge. Typically, not only do you need to know SQL for example, but also in which tables to look, how to find the data relevant to your test and what the data means. This very often limits this to developer and is an aspect not considered in later phases of testing where most of testing is performed. The simple fact is that the most important aspect of the application gets the least attention with testers and users focusing on the user interface. It is for this reason Original Software developed technology to simultaneously test the application's affect on the database at the same time as the user interface. Combined with a rules based approach, it means thorough examination and validation of the database continues through all phases of testing, building on knowledge and extermination otherwise hidden errors.

Finally, control of test data ensures that every test starts with a consistent data state, essential in maintaining your data in a predictable state at the end of the test. Without data consistency, and repeatability, testing will revert to a manual approach or complex algorithms to allow for changing data. By considering data to be an essential part of a regression test pack, it enables a much simpler approach to validation and the resulting improvement in productivity. The AQM effect is apparent even from initial implementation. The technology that enables a much wider view of the data under test is what is important in testing. The ability for the tester to simultaneously and clearly validate all aspects of the user interface and the associated database updates provides a repeatable way to carry out in-depth tests and deliver robust applications in a shorter timescale.

Testing on traceable live data is a thing of the past!

DATA TIP:

Create a test data privacy team composed of DBAs, a database security analyst, security officer, a data architect, and an enterprise architect who will oversee the test data privacy initiative.

Intelligent Database Management & Verification

TestBench from Original Software, enables you to test all aspects from Web or GUI down into the underlying system database. It sets the standard for quality in application testing and provides a unique solution for the data aspect of your quality process, designed to perform rigorous, deep and wide testing.

> “...a unique solution for the data aspect of your quality process, designed to perform rigorous, deep and wide testing.”

As already mentioned, using production data in a test environment can increase data security risks. Using ***TestBench's*** scrambling avoids this issue by masking or mixing data to make it untraceable so that valuable 'production-like' data can still be used and any legal penalty avoided.

The integrity of the database is key in most significant business applications, and it is also one of the hardest aspects to check. ***TestBench*** provides real in-depth and independent verification of the database showing developers exactly what happened; writes, updates, deletes file by file, record by record, field by field and program by program, whether batch or Interactive. It shows this wherever it happened, whether it was to be expected or not, and provides the user with the ability to create rules to check integrity and validity as the events happen. Testers are also able to see detailed effects in the database made as a result of a test without having to go searching for it. Setting pre-defined 'Data Rules' ensures that all database events comply with business or test-based rules. As these data rules accumulate with knowledge and time, the database aspects are subject to increasing examination, closing the net on illusive errors end ensuring accuracy where it counts, in the database.

When developing and testing systems, mistakes do happen, it is part of the process. But it can be frustrating if in a system test you have spent a couple of days getting some transactions to a critical stage, only to encounter a problem which invalidates or corrupts the data that you have carefully created. Do you carry on and compensate, repeat the process, get a restore done? With ***TestBench***, you can simply roll back quickly to a convenient checkpoint. Setting a checkpoint takes a few seconds and you can build them automatically into test steps so you have complete control.

> “Setting a checkpoint takes a few seconds and you can build them automatically into test steps so you have complete control.”

Take control of your data and avoid your own test data massacre.

Discover Application Quality Management at origsoft.com today!

About Original Software

With a world class record of innovation, Original Software offers a solution focused completely on the goal of effective quality management. By embracing the full spectrum of Application Quality Management across a wide range of applications and environments, the company partners with customers and helps to make quality a business imperative. The solution encompasses a quality management platform, manual testing, full test automation and test data management, all delivered with the control of business risk, cost, time and resources in mind.

More than 400 organisations operating in over 30 countries use Original Software solutions. Current users range from major multi-nationals to small software development shops, encompassing a wide range of industries, sectors and sizes. We are proud of our partnerships with the likes of Coca-Cola, Cargill, HSBC, Unilever, FedEx, Pfizer, DHL and many others.

European Headquarters
Basingstoke, UK
solutions.uk@origsoft.com
www.origsoft.com

North American Headquarters
Chicago, USA
solutions.na@origsoft.com
www.origsoft.com