

# Test Data Management

## Bloor: Independent Research Information

### How your IT department is breaking the data protection laws

By: [Philip Howard](#), *Director of Research – Technology*  
Copyright © 2006 Bloor Research



Now, here's a dirty little secret that most of the people that concern themselves with corporate governance and compliance don't know about.

Leaving aside very small print that may have been inserted into documents that no-one ever reads, the data protection laws mean that personal information may only be used for explicit purposes that support the reasons for which the information was gathered in the first place. For example, if I buy books from Amazon.com then Amazon is entitled to store that information, analyze my reading preferences and use the results to send me targeted offers. However, it is not, without my explicit consent, allowed to pass that information on to third parties for other purposes. Nor is it allowed to use that information in other ways or allows other people to view that data, even within Amazon.

Let's change the subject. Companies like Amazon are constantly developing new and/or improved applications to better serve both its customer base and its own internal purposes. This involves an often complex and lengthy development process. Further, it also involves the testing of those applications before they go into production, so that you can ensure that they work properly. Now, there are a variety of such tests. Some, for example, test the application logic: if I add "a" to "b" do I get "c" as I was expecting, or do I get "d"? Other tests, however, simulate a live environment, and this is where we get back to the data protection act.

If you are simulating a live environment, how does the IT department do that? The simple answer is that it grabs a bunch of data (technically, a sample) from its production database and runs that through the testing process. This isn't an issue if the application is concerned with stock control but it is an issue if the application deals with individuals, as in a consumer sales application or a human resources application. In these cases, unless you have formal consent from the people whose data you are using, then simply sampling the production system for test data is illegal: you are using the data for purposes for which it was not provided and you are making that information available to people not authorised to see it.



Okay, if you're a compliance officer, things aren't quite as bad they seem. Not quite as bad. Most IT departments know that this data can't be used (though I expect some ignore this). It therefore has to be "scrubbed". What this means is

that identifying characteristics have to be removed from the data so that it is legal to use it. However, the process of scrubbing is lengthy, complex and expensive—it can often involve many days of consultancy—as a result of which test data is often only partially scrubbed (which means that it is still illegal) at best.

I am pleased to say, however, that help is at hand. Grid-tools has developed a product called Datamaker that generates anonymised test data from the production database (as opposed to the use of test scripts, which is the traditional approach) in order to populate the test environment. You won't find anything about the product on the company's website, or any of its other products if it comes to that, but I can tell you that the product is currently installed in its first customer site, which is inside a major UK government department that cannot afford to be seen to flout the data protection laws. If you have similar concerns then you could do worse than contact Grid-tools to find out more.