Test Data Privacy

Protecting Your Most Important Asset: Information
Start Your Autonomous Digital Enterprise Journey with a Mainframe-Inclusive DevOps Toolchain

Compuware joins BMC to empower the next generation of developers to mainstream the mainframe. With unparalleled mainframe operations management and agile application development and delivery, BMC Compuware provides a mainframe-inclusive DevOps toolchain that increases volume and velocity.

Bring apps that run on the mainframe to market faster with increased quality throughout the process of analyzing, building, testing, deploying, monitoring, and tuning apps and services. Our combined offering accelerates each client’s evolution to an Autonomous Digital Enterprise.
Our Information is Everywhere

Information is one of the most valuable assets of modern life. Information and Privacy Commissioner of Ontario, Canada Ann Cavoukian framed it best when she said that data is the new lifeblood of the new economy.

In recent years, the collection, processing, and commercialization of data has increased dramatically. Every day within the EU, organizations, government agencies, and individuals transfer vast amounts of personal data across borders.

More and more, individuals are sharing their personal information without being fully aware of the risks involved, knowing where their data is held, or by whom. Whether they are amenable to it or not, their electronic footprint expands with every piece of connected technology they use. Although much of this data collection translates to better benefits in terms of more customized services and offerings for EU citizens, the more eyes on their personal data, the greater the chance their private information will be compromised.

Trying to address this situation is made more complex by the fact that many businesses are trading across borders. Existing national data protection laws are no longer enforceable in this cross-border, computerized, and highly interconnected world.

“People feel they have little to no control about how their data may be collected by online companies.”

Microsoft Data Privacy Day Privacy Survey, Jan. 2013

“Data is the new lifeblood of the new economy.”

Ann Cavoukian
Information and Privacy Commissioner of Ontario, Canada
The European Union (EU) has taken this challenge seriously and is in the process of reforming the 1995 Data Protection Directive to further protect the personal data of its citizens. The proposed new rules provide an updated legal basis for a modernized and comprehensive approach to data protection, giving citizens more control over their personal information while ensuring free movement of personal data across the EU.

In March 2014, the European Parliament voted on the European Commission’s data protection reform which is represented as two important pieces of legislation affecting the personal data of EU citizens (both in draft):

- **The EU Data Protection Regulation** – establishes a general EU framework for personal data protection and the free movement of such data

- **The EU Data Protection Directive** – covers personal data processed by competent authorities with the purpose of prevention, investigation, detection, or criminal prosecution of criminal offenses and the free movement of data for these purposes

The objectives of the reform include:

1. **Simplifying the regulatory environment.** Currently, data controllers may have to deal with as many as 28 different national laws and requirements within the EU. Under the new rules it is proposed that companies will deal with a single law and be answerable to one data protection authority no matter how many countries they do business in.

2. **Providing individuals control over their personal data, particularly electronic data.** Due to the ease with which data is collected electronically individuals have lost control over their own data, not knowing by whom it is collected, and what disclosures are being made to other parties. The new law will guarantee data protection while ensuring the free flow of data.

3. **Making it easier for police and judicial bodies to investigate and prosecute in criminal matters.** Under the current directive, there are legal uncertainties for both individuals and law enforcement authorities, as well as practical difficulties of implementation. The data protection reform will establish one set of data protection rules and ensure they are applied uniformly across the EU.

4. **Increased responsibility for parties processing data on behalf of others (data processors).** The new regulation will mean companies processing data on behalf of other companies will have the same level of responsibility as the company they are working on behalf of.

### Analysis of Primary Causes of Data Breach

- **Negligence**: 24%
- **Malicious or criminal attack**: 37%
- **System glitch**: 34%
What EU Citizens Can Expect

It’s estimated that 250 million citizens use the Internet in Europe on a daily basis. Whether engaging with others on a social networking site or supplying contact information to online vendors, individuals are sharing their private information more now than ever before. The reform is intended to give people more control over their private information through stronger data protection rules that provide:

1. **The ‘right to be forgotten.’** When people no longer want their data to be processed and there are no legitimate grounds for retaining it, the data will be deleted.

2. **More focus on protection.** Companies in certain circumstances will need to appoint a Data Protection Officer and also conduct assessment of the impact of new systems or changes to existing systems handling personal data.

3. **Faster notification.** The reform requires organizations to notify the commissioners and the citizens within a reasonable time period if their information is accidentally or unlawfully destroyed, lost, altered, accessed by, or disclosed to unauthorized persons and the breach will lead to damage or distress of the citizens.

What Companies Can Expect - The Highlights

Technology and globalization have profoundly changed the way information is collected, processed, and used. Companies will now be required to take additional measures to protect citizens’ private information or face serious censures and fines.

The following are some of the key changes organizations can expect to take effect:

- Companies that process over 500 personal data records a year or have over 250 employees will be required to have a Data Protection Officer.

- Organizations will only have to deal with a single national data protection authority — in the EU country where they have their main establishment.

- In the event of a serious data breach, organizations will be required to notify the data protection authority as well as the impacted citizen/citizens ‘without undue delay.’

- Data processors (companies processing data on behalf of company/companies) are obligated to report a breach to the data controllers of the company where original data resides.

These changes will not only clarify organizational responsibility through the roles and responsibilities, but they will also define some operational system compliance requirements for businesses.

These rules, of course, translate into the requirements for IT System consideration.
Secondary Data Usage With Personal Data - Required?

Organizations will be expected to adopt a higher level of control and protection of their data, called “Data Privacy by Design” and provide proof of these practices under an information audit. Privacy-enhancing technologies have evolved over the past decade and data controllers and processors within organizations will be expected to have performed due diligence in their use.

Many IT departments when considering application system testing state that they require ‘as is’ production data against which to test and that personal data needs to be intact. But this is just not the case. A number of arguments can be posed against this statement; however the main point is that real personal data can be masked and is therefore not required for application testing. In a practice called Static Data Masking (SDM), data is extracted from a production environment and copied into a test environment, and personally identifiable information can be masked before testing.

Is Your Data Safe?

- 80 percent of respondents were willing to share personal information if the company lets them know upfront how they are going to use it*
- It is often unclear as to who is responsible for the security of customer data within organizations
- For many organizations, live customer data files represent an easy and cheap source of data to use when testing applications
- Practices that relate to all uses of live data must be evaluated to assess risk, and safeguards implemented to ensure data security

* Attitudes on Data Protection and Electronics Identity in the European Union, June 2011
Static Data Masking

The concept of SDM is defined by Gartner (ref) and involves data technology that mitigates misuse of production data (primary data usage) in non-production environments such as test, analytics, or training through the masking or desensitization of the data prior to usage in these environments.

When using this technique there are a few fundamental factors to consider:

1. The integrity of business processing/testing conditions are maintained. This dictates that the masking of sensitive data must be done in the context of the business logic associated with the sensitive data, otherwise the application will reject the data, and that the masked data should drive the application logic at run time to meet the desired test conditions.

2. Masking or desensitization is applied consistently across data structure and platform. This is in effect saying where a sensitive data item is being masked, for example account number 12345678 is masked to become 74925910, the change has to be propagated across all the data structures that contain that account number irrespective of the platform or platforms that data resides on.

If a data privacy solution is embraced that considers and addresses these factors then production data can indeed be successfully desensitized, hence be compliant, while being used for secondary data purposes such as testing.

The Plan to Address the EU Directive

The European Commission has stated that the data protection reform is ‘now irreversible’ and that even though the path to implement the text still needs to be agreed upon by both Parliament and the Council before it becomes law, they are well on the road to a law being in place. But are European businesses prepared to address this new reality?

A global CIO survey commissioned by BMC Compuware and conducted by Vanson Bourne revealed several interesting findings on this subject:

- 20 percent of companies do not mask or protect their customer data before providing it to outsourcers for application testing purposes
- 43 percent of respondents that share customer data do not understand current data protection laws and regulations
- 87 percent of organizations that do not mask customer data before passing it to a third party rely on Non-Disclosure Agreements (NDAs) to protect their customer’s data

These results set off IT alarm bells not only in light of the new EU directive that is coming but also against a backdrop of existing regulation. Let us not lose sight of the fact that since 1995 when the EU issued the Data Protection Directive 95/46/EC that there are principles governing the usage of personal data that need to be complied with at present. The new EU laws will encompass and supersede these. Some of the existing principles relevant to this paper are:
1. Personal data shall be obtained for lawful purposes and shall not be processed in any manner incompatible with those purposes.

2. Personal data shall be adequate, relevant, and not excessive in relation to the purposes for which processed.

3. Personal data shall be accurate and up to date and not kept for longer than is necessary.

4. Appropriate technical and organizational measures shall be taken against unauthorized or unlawful processing and against accidental loss, destruction, or damage to personal data.

5. Personal data is not transferred outside the EEA without adequate protection.

It is evident that companies are compliant when looking at the fundamental elements of data privacy based on existing legislation, never mind the new regulations.

When it comes to testing IT systems there are two factors to consider. Firstly, that testing needs to be thorough in order to maximize the benefits to users and minimize the potential damage and distress that such systems cause. Secondly, the usage of live personal data in systems testing creates risk in that such systems are less secure and challenges data protection compliance.

“62% of theft is carried out in IT or customer service.”

The Enemy Within – 2010 Imperva Survey

Examples of Personally Identifiable Information

- Name
- Address
- Telephone Number
- Email Address
- NI Number
- Account Number
- Customer Number
- Credit Card Number
- Debit Card Number
- Bank Details
- Employment Details
- Family Details
- Medical Information
Test Systems - Creating Low Risk

It is important that organizations implement systems that minimize risk of data protection principle breach and that the decision process is appropriate for the level of technology available, considers the cost of implementing a solution relative to their sector and size, and lastly considers the potential impact or unwanted disclosure of personal information and particularly sensitive information. These are the factors that the EU commissioners will consider when determining the seriousness of the breach if it occurs and therefore the size of the monetary penalty that they issue.

It is in the best interest of the organization to consider preventative measures to protect personal data, especially when using such data for system testing. With the state of technology and software available in today’s world it would not be an advisable strategy to presume that live data identifying individuals could be used in its original form for testing. So, what are the recommended approaches to this?

Well, there are a few approaches that can be taken such as:

1. Generating dummy data in test: In the case of complex systems or systems with large volumes of data required for performance testing, this in reality is not a feasible option to pursue.

2. Anonymization: This is a highly effective approach and it also has organizational benefits if used in a Static Data Masking context where data optimization can also be achieved while increasing quality and turnaround of testing cycles.

The usage of live data in test systems is described by the Information Commissioners Office (ICO) in the UK as something that should be avoided. Indeed the advice is that if there is no practical alternative that alternative methods of testing should be sought rather than using live data. For IT systems alternative methods of testing cannot be sought, however, as multiple anonymization packages exist this is highlighted as the preferred approach due to its minimization of risk.

It is common practice in the IT industry that live data is simply copied into test environments. Taking this approach as it is often encountered would not be taken favorably when there is a breach and would result in increased monetary penalties as it infringes data protection principles and therefore is an unlawful practice.

Organizations must remember the consideration of personal data being used for lawful purposes and that they need to notify citizens of their purpose for collection of data, especially when the purposes are not obvious to a customer. Such as is the case with using personal data for system testing.
Where to Start?

Addressing the data protection directive requirements is daunting and many organizations are at a loss as to where to start. This is a project that has a very wide reach — it affects all applications, all data repositories, and all users across the IT organization as well as the business. There are high risks, high stakes and, potentially, high costs.

Many information security teams have mobilized and are engaging with IT development and testing teams to audit what is done today and determine what they will need to do to be compliant. What they are uncovering is an unnerving picture. Due to the proliferation of live data and the extent to which it is used outside the production environment, the sheer size of the undertaking to become compliant is something they were not prepared for. Many organizations are admitting that they need help and are seeking partners to guide them through this process. Their advice to the wider business community is, ‘Don’t delay, act now, it may take you longer than you think’.

It is important to review your security strategy and be clear about what is required for compliance. You will also need to know how and where data is used in your organization today before you start developing your project plan. There are detailed activities that you may not even have thought about, so planning resources, tooling, and the selection and implementation process needs to take the highest priority.

There are also financial implications to consider. The EU Data Protection Directive changes are expected to give rise to additional compliance costs. There are some preliminary numbers available surrounding the annual costs to execute the law — the UK government has, for example, estimated that the annual cost of the proposed measures will be between £100 to £360 million — but there’s been little discussion about the costs of initial implementation of the compliance framework.

The good news is, this cost can be influenced and minimized by ensuring the right experience, methodology, and tools are used during the all-important implementation and deployment phases of the compliance project. Read more about the 5 key steps to attain, prove, and sustain compliance by going to compuware.com/privacycompliance.

BMC Compuware has been engaged with many organizations across the globe, helping deliver successful, large-scale data privacy projects. The following is BMC Compuware’s Test Data Privacy approach.

“Data Breach in the proposed Directive could mean penalties of up to €100m or five percent of the organization’s global annual turnover.”

BMC Compuware’s proven process underpinned by expertise and market-leading technology.
BMC Compuware’s Solution: Protect Your Customer’s Data and Your Business

BMC Compuware’s Test Data Privacy solution provides a proven means to mitigate the risk of data breaches through static data masking (SDM) in complex IT environments when dealing with secondary data usage.

BMC Compuware’s Test Data Privacy offering includes an end-to-end combination of technology, expertise, and best practices needed to support data protection initiatives across platforms and, thus, the enterprise. The BMC Compuware solution provides an automated process for sub-setting and desensitizing copies of primary data containing PII (Personally Identifiable Information). This data may be used for testing, QA, or possibly for transmission to other business partners.

Technology

BMC Compuware has been working with customers to desensitize data outside of primary data usage for many years and the toolset has evolved based on experience from large scale project requirements. As with any project, sustainability and maintainability as well as cost are key considerations to be addressed.

BMC Compuware initially started with an object and field targeted approach where masking is defined against named fields and columns in the tables/files to be masked. However, this was not a sustainable approach for the level of complexity required on the mainframe, and between mainframe and distributed environments. The answer was to instead associate masking with the types of sensitive data elements such as card numbers, account numbers, names independent of object (file/table), or platform. BMC Compuware refers to this technique as Dynamic Privacy Rules.
**Data Privacy Manager and Dynamic Privacy Rules**

BMC Compuware’s enterprise-wide data privacy manager enables masking requirements, often defined by the business, to be applied consistently, even when data spans both mainframe and distributed platforms, different databases, and a variety of file types. A unique Data Element approach ensures high efficiency and time saving by creating masking rules against categories of data instead of individual fields/columns. Typically, organizations have hundreds, if not thousands, of columns/fields that contain sensitive data. However, there are rarely more than 10-15 categories. BMC Compuware’s approach therefore greatly reduces the number of necessary rules, typically from hundreds or thousands to just a handful.

At run time, BMC Compuware’s data masking uses technology referred to as the Common Disguise Engine which performs data normalization into and out of the masking process, which ensures consistent masking across files, databases, irrespective of the operational platforms or encodings.

**File-AID**

BMC Compuware’s market-leading File-AID technology underpins test data privacy with several key differentiators:

1. Data selection and sub-setting capabilities allow focused and relevant data to be extracted while maintaining the integrity of the data; an essential requirement for high-quality test data.

2. Diverse mechanisms for data obfuscation; format preserving encryption and translation (a.k.a. “look-up”) being the two most commonly used.

3. Powerful data load capabilities to refresh, update, or append target environments.

4. Integrated tools to allow developers to discover and maintain data file and application relationships.

5. Providing functionality to Transform, Move, and integrate.

This technology is designed to deliver an efficient, repeatable process ensuring a sustainable mechanism for the long-term success of your project. With the BMC Compuware File-AID products at the core, organizations are assured of consistent data disguise techniques being applied across all environments. These products provide a unique solution to help companies address requirements to protect personally identifiable information throughout the application testing process.

Together, Topaz Workbench and File-AID enable developers to disguise data consistently for each project by entering a rule only once. So for each project they can then quickly and easily find, create, extract, transfer, convert, load, and edit protected data.
Best Practices

Undertaking a test data privacy project is a significant effort for any organization; it is recognized that software tools alone are not the answer. BMC Compuware’s proven best practices have been built up over many years of successfully delivering projects for our clients. They provide the guiding principles to build or enhance the process of preparing secure data environments to support the IT functions related to the usage and delivery of data in production-like conditions.

These best practices address the diverse technical, organizational, and political challenges associated with the protection of data. Built around a project-based methodology, they also offer insight on proven techniques, reference materials, technical know-how, etc. The methodology outlines the phases, activities, tasks, and deliverables that encompass the implementation of a complete data privacy solution within the context of an existing quality assurance lifecycle.

Addressing Compliance

Organizations are starting to realize that “do nothing” is no longer an option when it comes to data protection regulation. Getting a clear understanding of the compliance requirements and designing an accurate road map of how to get there is a fundamental requirement for doing business in the future.

The common IT practice of using live data for testing will no longer be acceptable and organizations will need to demonstrate due diligence in their efforts to protect information about their customers.

Organizations that have adopted BMC Compuware’s solution have identified many benefits that have resulted from their data privacy initiatives, including:

1. Sub-setting of data has meant that less data is required in the test environments, resulting in reduced costs for both storage media and CPU power.
2. Testing cycles are much faster due to the fact that tests are using reduced, more focused test data.
3. The quality of testing has improved as testers work with higher quality data.
4. IT is able to deliver more software releases to the business, increasing competitive agility and realizing greater customer satisfaction.

For more information, please contact your BMC Compuware Account Manager.

*These statistics were taken from a survey administered to 350 CIOs at large companies covering a cross-section of vertical markets in the U.K., France, Germany, the Benelux, Italy, Japan, Australia and the U.S.