



Data Security in Development & Testing

Sponsored by
Micro Focus

Independently conducted by Ponemon Institute LLC

Publication Date: July 31, 2009

Data Security in Development & Testing

Findings from a survey of IT practitioners in the UK and US

Dr. Larry Ponemon, July 31, 2009

I. Introduction

An overlooked threat to the safeguarding of an organization's sensitive and confidential data is the vulnerability of personal and business information used for testing and application development. Organizations may think their test data is immune from data security threats because testing occurs in a non-production environment. However, test environments are less secure because data is exposed to a variety of unauthorized sources, including in-house testing staff, consultants, partners, and offshore development personnel.

According to this study, an organization's data is at risk primarily due to internal threats. The three most common reasons for data breaches amongst survey respondents were negligent insiders, malicious insiders and third party flubs. The following three incidents in the United States and the United Kingdom support this finding.

This first incident involves an ex-Goldman Sachs programmer who allegedly transferred the firm's proprietary trading code worth millions of dollars to a computer server in Germany. If the code is disseminated, others may have access to it as well. In the second incident, HSBC was fined £3 million by the Financial Services Authority in the UK for a failure to adequately protect customers' confidential details. The company lost large amounts of unencrypted customer details when it was sent to third parties. The third incident was also in the UK and concerns insider negligence. In this case, the personal details of 25 million British citizens were lost when two computer discs were being sent by internal mail from one department to another.

The purpose of this study is to learn how vulnerable real data used in the development and testing environment is to a potential data breach. We surveyed 701 US and 652 UK practitioners involved in systems development and testing. Findings indicate that these practitioners are not confident of their organizations' ability to safeguard data used in testing and application development.

Participants in our study were asked about the following issues:

- What are the types of real data used in testing?
- Does the organization use the same safeguards for test data as it uses for other sensitive and confidential information it collects, uses and stores?
- What data security measures are in place to protect test data?
- Who is responsible for protecting test data?

II. Summary of most salient findings

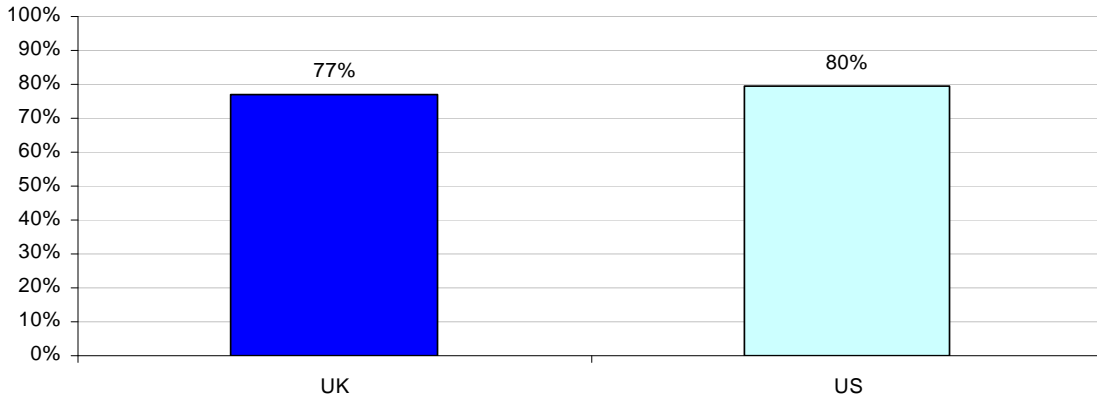
Following are the most salient findings of this survey research. Please note that most of the results are displayed in bar chart format. The actual data utilized in each figure and referenced in the paper can be found in the percentage frequency tables attached as the Appendix to this paper.

Real production data is used frequently in application development and testing.

The majority of respondents say they use real data in application development and testing. According to Bar Chart 1, 80% of respondents in the US and 77% in the UK report that they use real production data as part of their application development and testing process. The most common data types used in both countries are similar. Further, 63% of UK respondents and 61% of US respondents report that they refresh real data for application testing either every time or

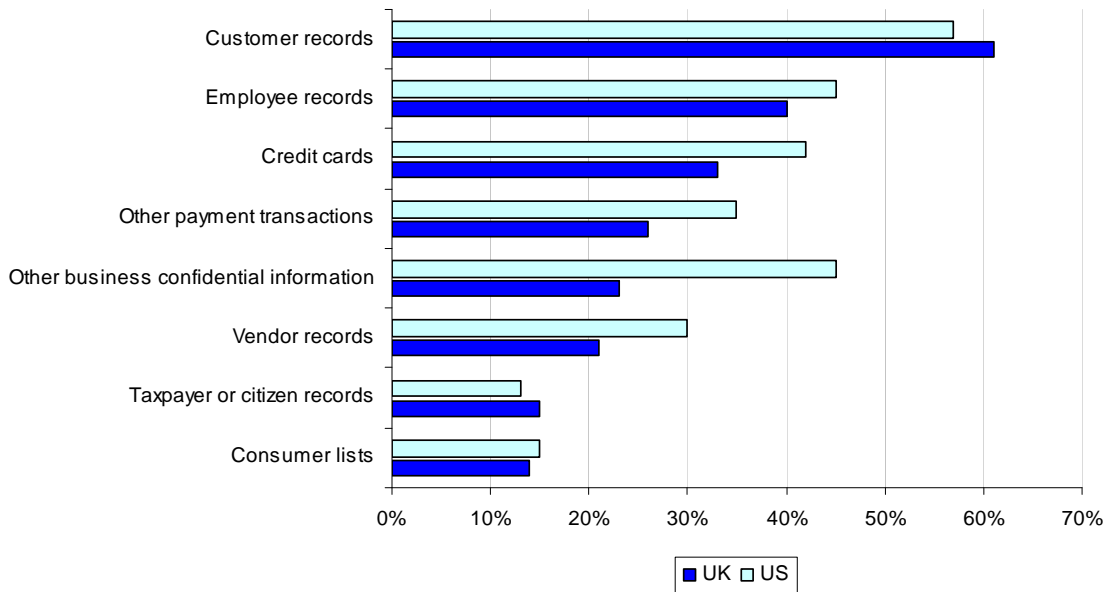
frequently (see Question 8b). Hence, this creates the possibility of multiple insecure copies of large files containing sensitive or confidential data.

Bar Chart 1
Do you use real production data as part of your application development and testing process?
Percentage Yes response.



As revealed in our study, highly sensitive data is being used for development and testing. As shown in Bar Chart 2, customer records, employee records and credit cards are the most frequently used data.

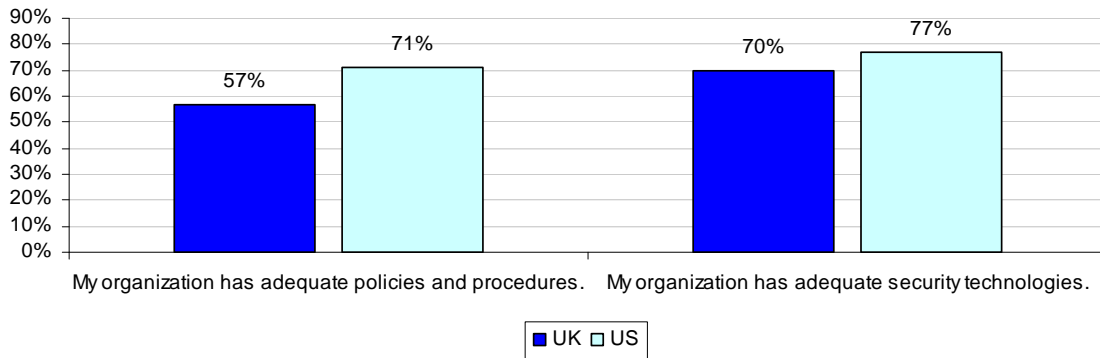
Bar Chart 2
Types of real data used for development and testing



The majority of respondents hold the belief that their organizations do not have adequate policies and technologies in place to protect real data used in development and testing.

Bar Chart 3 shows respondents in both the UK and US are less than enthusiastic about their organization's ability to secure real data used in application development and testing. Based on a five-point adjective scale (from strongly agree to strongly disagree), 57% of UK and 71% of US respondents strongly disagree, disagree or are unsure that their organizations have adequate policies and procedures. Seventy percent of UK and 77% of US respondents strongly disagree, disagree or are unsure that their organizations have adequate security technologies.

Bar Chart 3
Strongly disagree, disagree and unsure responses (combined) for two questions about the respondent organizations' ability to secure real data used in development and testing.

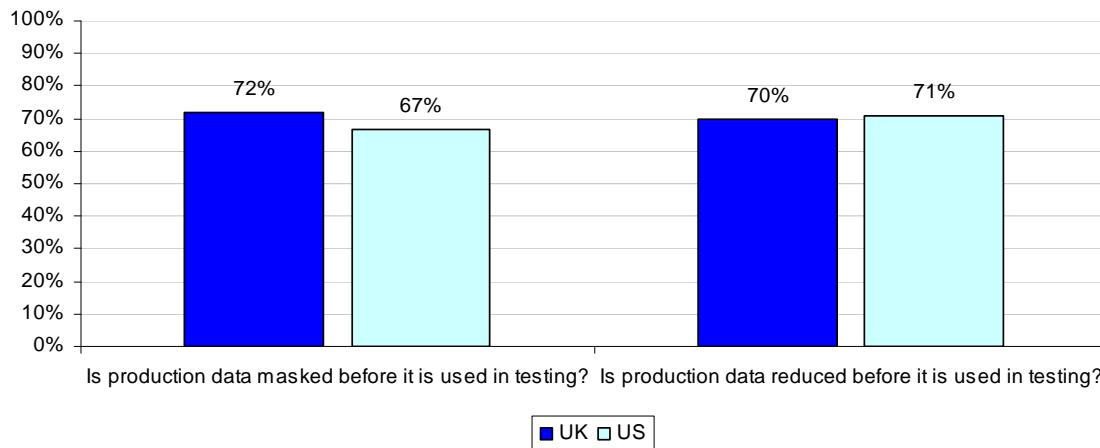


Respondents in the UK appear to hold a more positive perception about the adequacy of their organizations' policies, procedures and security technologies available for protecting real data used in development and testing.

Most respondents in this study say their organizations do not mask or reduce (subset) production data before it is used in development and testing.

As shown in Bar Chart 4, 72% of UK respondents and 67% of US respondents say they do not use masking methods to protect production data in development and testing. Similarly, 70% of UK respondents and 71% of US respondents do not reduce or subset production data before using it in development or testing.

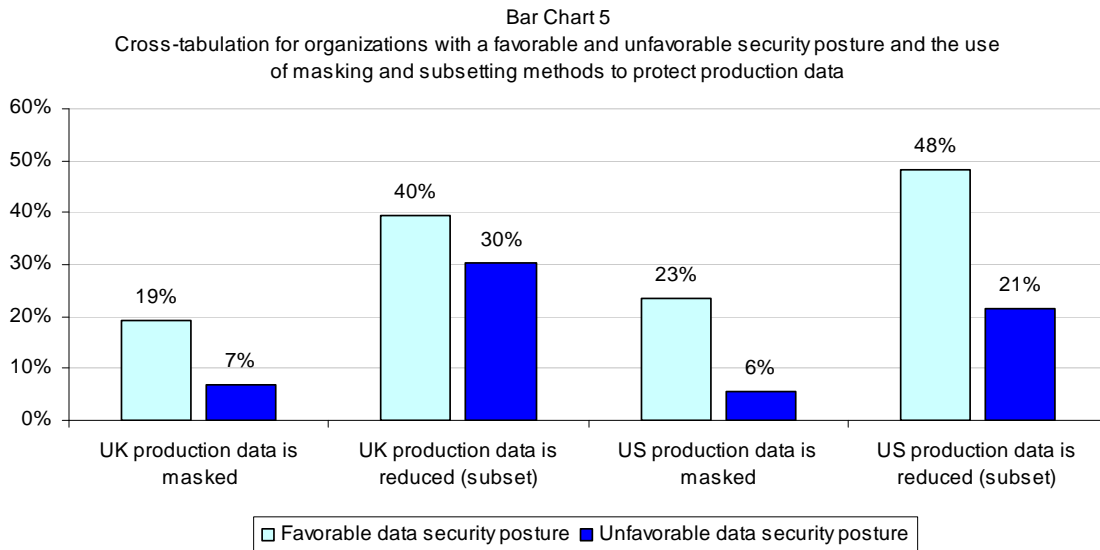
Bar Chart 4
Respondents who say they do not mask or reduce real data in development and testing



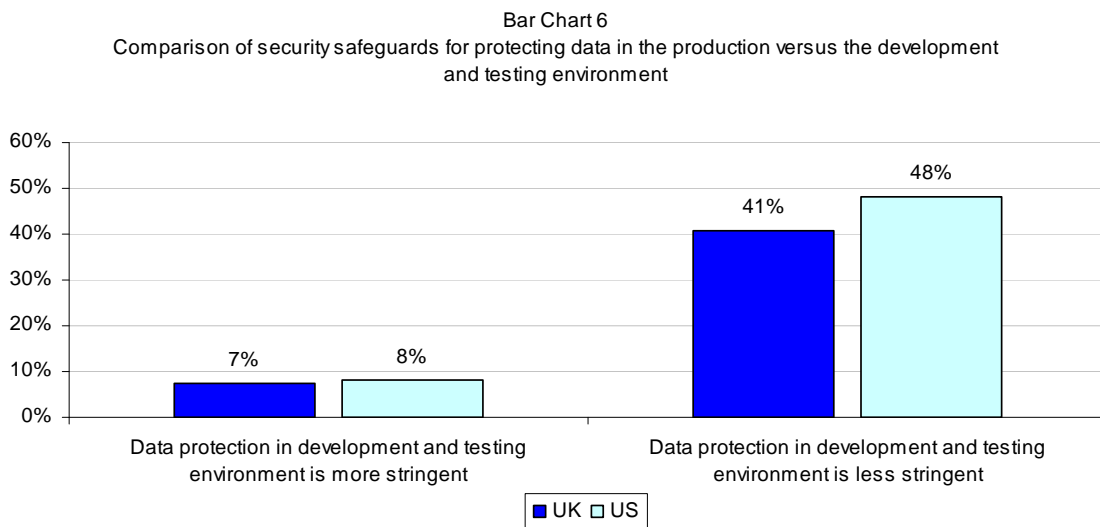
Organizations that have a favorable data security posture are much more likely to mask or subset production data before it is used in development and testing.

Five attribution questions (see Question 1) were used to measure respondents' perceptions about their organization's security posture with respect to real data used for development and testing. Respondents who chose "strongly agree" or "agree" to all five attributions were defined as having favorable perceptions about their organization's security. Those who did not choose "strongly agree" or "agree" to at least one attribution were defined as having an unfavorable perception.

Bar Chart 5 clearly shows in both the UK and US that organizations with a favorable security posture are much more likely to utilize masking and reduction techniques to secure real data used in development and testing. Ponemon Institute believes masking and reduction methods are among best practices for companies seeking to protect data in development and testing.



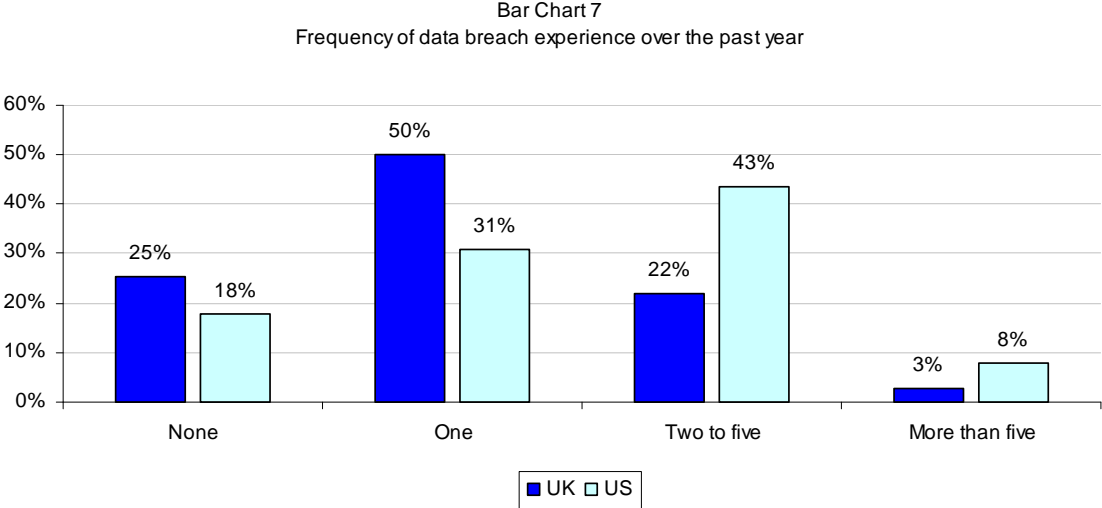
The majority of respondents acknowledge that data protection in the development and testing environment is less stringent than in production.



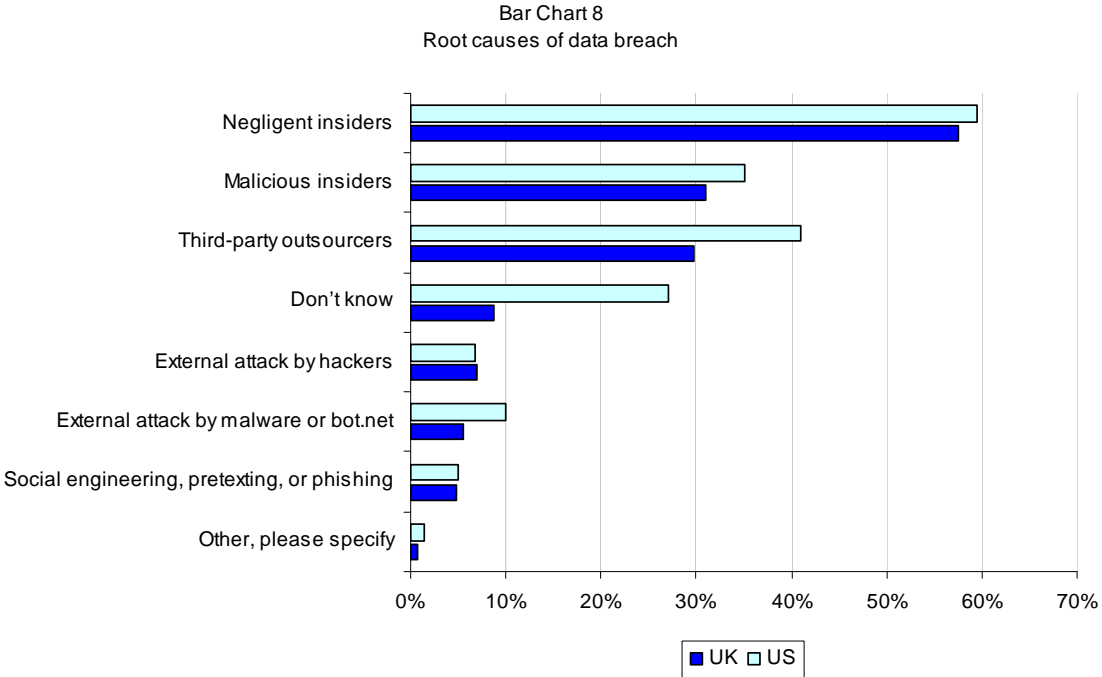
As shown above in Bar Chart 6, 41% of UK respondents and 48% of US respondents believe the protection of real data in the development and testing environment is less stringent than comparable safeguards over real data in production.

The risk that data used for development and testing purposes will be lost or stolen is real.

As shown in Bar Chart 7, 75% of respondents in the UK and 82% of respondents in the US report that their organizations have had one or more data breaches.

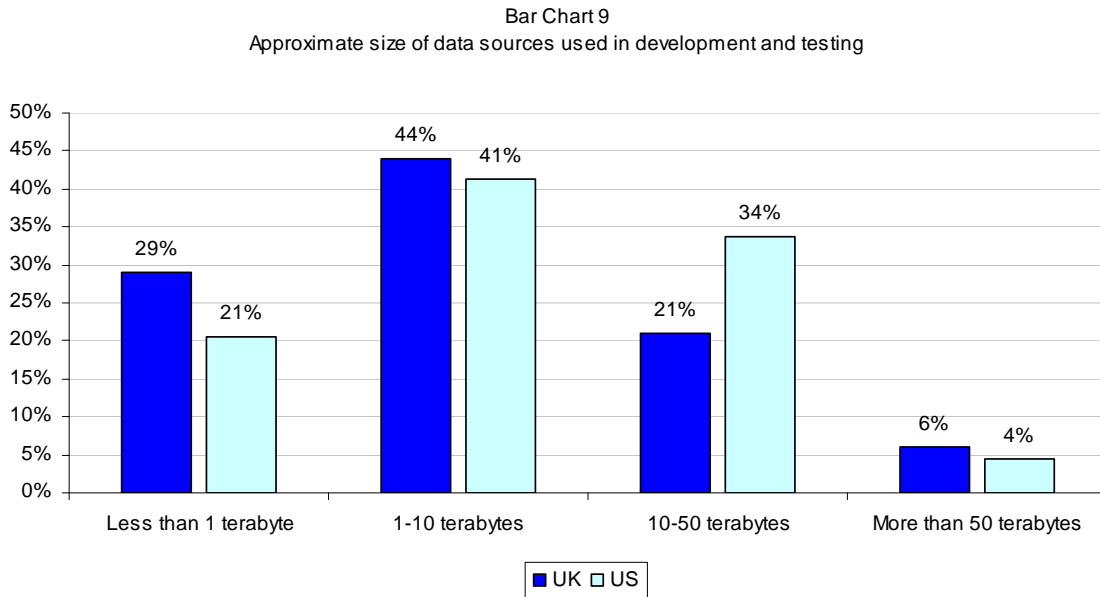


Bar Chart 8 shows the root causes for data breach. As is evident from Bar Chart 7, most organizations are experiencing data breaches. However, based on the study’s findings there does not seem to be proper precautions in place to protect data in testing and development from the root cause of a data breach—the insider threat. The top three root causes in both the UK and US are negligent insiders, malicious insiders and third-party outsourcers.



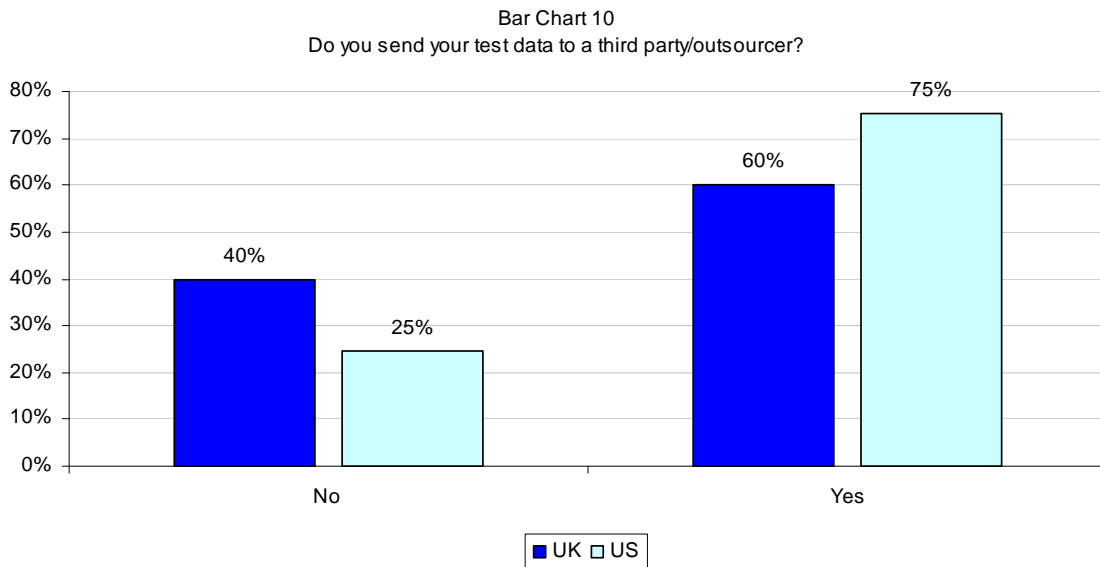
Large amounts of production data are used in development and testing

According to Bar Chart 9, as many as 71% of respondents in the UK and 79% of respondents in the US say they use files with more than one terabyte of real data in development and testing.



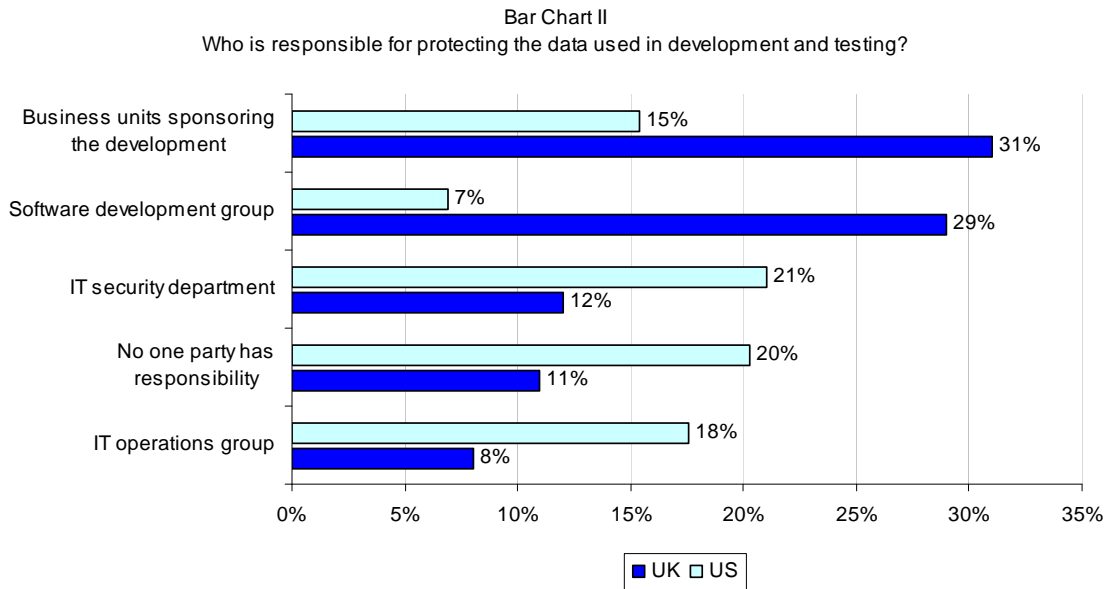
Respondents admit that real data is frequently shared with offshore outsourcers and other third parties, thus creating additional privacy and data protection risks.

Bar Chart 10 shows that 75% of US respondents and 60% of UK respondents send real data to third-party organizations for development and testing. This finding, coupled with the finding that third-parties was one of the most frequently cited root causes of a data breach, suggests these organizations face the risk of third-party mistakes or flubs.



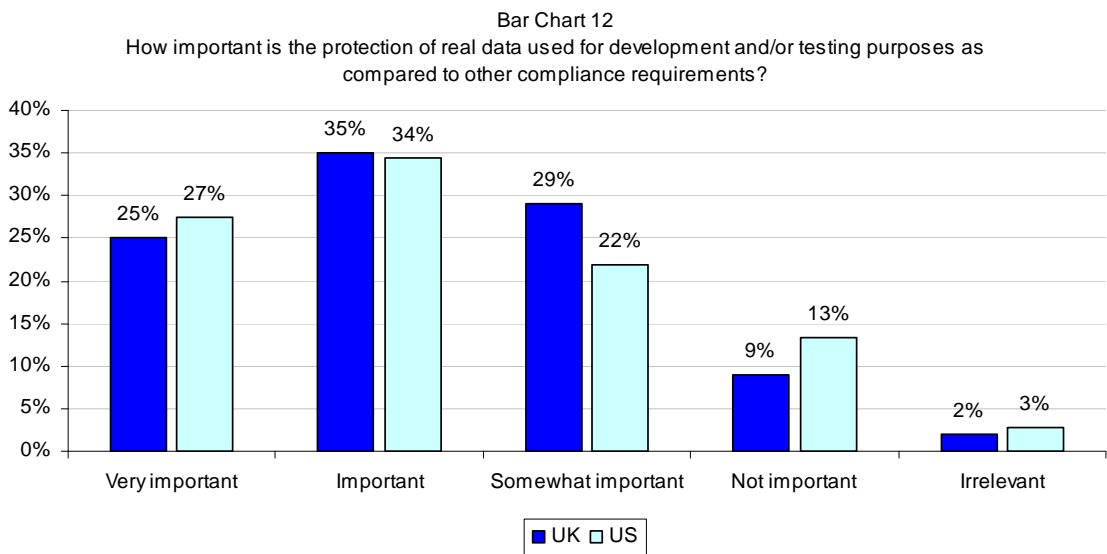
Respondents do not see one organizational function as owning responsibility for data protection in the development and testing environment.

Bar Chart 11 shows marked differences between responses of UK and US respondents. As can be seen, 31% of UK respondents, as compared with only 15% of US respondents, believe business units that sponsor development are responsible for protecting real data. In contrast, 21% of US respondents, as compared with only 12% of UK respondents, believe the IT security department is responsible for protection of real data used in development and testing.



Respondents overwhelmingly believe that the protection of real data is important to achieving IT compliance within their organizations.

Bar Chart 12 shows 60% of UK respondents and 61% of US respondents believe the protection of real data in development and testing is an important or very important compliance objective.



III. Caveats

There are inherent limitations to survey research that need to be carefully considered before drawing inferences from findings. The following items are specific limitations that are germane to most -based surveys.

- Non-response bias: The current findings are based on a sample of survey returns. We sent surveys to a representative sample of individuals, resulting in a large number of usable returned responses. Despite non-response tests, it is always possible that individuals who did not participate are substantially different in terms of underlying beliefs from those who completed the instrument.
- Sampling-frame bias: The accuracy is based on contact information and the degree to which the list is representative of UK and US IT practitioners. We also acknowledge bias caused by compensating subjects to complete this research within a holdout period. Finally, because we used a -based collection method, it is possible that non-responses by mailed survey or telephone call would result in a different pattern of findings.
- Self-reported results: The quality of survey research is based on the integrity of confidential responses received from subjects. While certain checks and balances can be incorporated into the survey process, there is always the possibility that a subject did not provide a truthful response.

IV. Methods

Table1 reports our sampling plan. Our sampling frame focused only on IT practitioners involved in the applications development and testing arena. In total, 11,009 individuals in the UK and 12,260 individuals in the US were invited to participate for nominal compensation. This resulted in a total sample size before screening of 652 (5.9% response rate) and 701 (5.7% response rate) in the UK and US, respectively.

Table 1 Description of the sample	UK	US
Total sample frame over eight (8) days	11,009	12,260
Bounce-back	981	1,013
Rejections on reliability	45	53
Total sample (before screening)	652	701
Response rate	5.9%	5.7%

Table 2 reports the respondent's organizational level. The majority of respondents are at or below the manager level within their organization. The technician level represents the largest segment.

Table 2 What organizational level best describes your current position?	UK	US
Senior Executive	0%	0%
Vice President	0%	1%
Director	8%	12%
Manager	23%	23%
Supervisor	21%	18%
Technical	35%	32%
Associate or staff	3%	5%
Administrative	0%	1%
Contractor	9%	8%
Total	100%	100%

Table 3 provides the full-time equivalent headcount as a surrogate for organizational size. As can be seen, the majority of respondents are employed by larger-sized companies with more than 5,000 people.

Table 3 What is the worldwide headcount of your organization?	UK	US
Less than 500 people	7%	4%
500 to 1,000 people	8%	9%
1,001 to 5,000 people	27%	20%
5,001 to 25,000 people	30%	23%
25,001 to 75,000 people	21%	30%
More than 75,000 people	7%	14%
Total	100%	100%

Table 4 reports the industry segments represented in this study. As shown, the largest segments include financial services, government, technology and software and healthcare.

Table 4 Industry segments	UK	US
Financial services	20%	23%
Government	15%	14%
Technology & Software	14%	18%
Healthcare & pharma	9%	10%
Retail	7%	3%
Services	6%	5%
Communication services	5%	6%
Other	5%	6%
Education	5%	4%
Transportation	4%	4%
Manufacturing	3%	3%
Hospitality	3%	3%
Entertainment & media	3%	3%
Total	100%	100%

The next set of data provides further information about employers. Tables 5a and 5b report on the companies listing and ranking on a major exchange (FTSE for UK and Fortune for US).

Table 5a Is your company publicly traded?	UK	Table 5b Is your company publicly traded?	US
Yes, FTSE 100	21%	Yes, Fortune 100	12%
Yes, FTSE 500	32%	Yes, Fortune 500	24%
Yes, not ranked in the FTSE	15%	Yes, not ranked by Fortune	21%
No	32%	No	42%
Total	100%	Total	100%

V. Concluding thoughts

Sensitive and confidential data is at risk in the testing and development process. The most likely threats to this data based on research and actual cases are malicious insiders, negligent insiders and third parties. Further, respondents in our study do not believe that their organizations have a strong security posture to address these threats. Specifically, they are less likely to agree that their organizations have the resources, technologies and motivation to protect the privacy rights of consumers, customers, employees and other individuals whose data is used for software development and testing.

Based on these findings, we encourage organizations to take the following steps.

- Understand that data can be breached when it is used for testing and development purposes. Based on Ponemon Institute's Cost of a Data Breach Study, each record that is lost or stolen could cost the organization an average of \$202.¹ Further, such a breach would also diminish the brand and reputation of the organization. This should help make the case that resources are needed to address this risk.
- Create a culture that empowers and rewards employees who respect and are accountable for protecting sensitive and confidential information. The importance of protecting privacy rights should be communicated from the top down.
- Ensure that third-parties and other vendors with access to the organization's sensitive data have data protection procedures and policies in-place. We also recommend a period assessment of the effectiveness of these safeguards.
- Finally, assess the risk to real data in the testing and development environment. Based on this assessment, consider implementing not only the appropriate technologies that will mitigate the risk but policies and procedures to safeguard data. The assessment and implementation should be part of a data protection strategy for the use of data in testing and application development.

¹ See Fourth Annual Cost of Data Breach Study, Ponemon Institute, January 31, 2009.

Appendix 1: Survey Results

The following tables summarize the results of an independent study completed by Ponemon Institute and sponsored by Micro Focus. Please note that all findings have been audited by Ponemon Institute's quality assurance. Fieldwork commenced on 8 July 2009 and concluded on 16 July 2009.

Description of the sample	UK	US
Total sample frame over eight (8) days	11,009	12,260
Bounce-back	981	1,013
Rejections on reliability	45	53
Total sample (before screening)	652	701
Response rate	5.9%	5.7%

S1. Do you use real production data (cloned, copies, etc) as part of your application development and testing process?	UK	US
Yes	503	558
No (stop)	149	143
Total	652	701
Percentage Yes	77%	80%

Sample size after screening question	503	528
--------------------------------------	-----	-----

Q1. Please respond to each statement about your organization using this five-point scale to express your opinion: 1=Strongly agree, 2=Agree, 3=Unsure, 4=Disagree, 5=Strongly disagree	UK 1 & 2 combined	US 1 & 2 combined
Q1a. My organization has adequate policies and procedures to protect real data used for software development and testing.	43%	29%
Q1b. My organization has adequate security technologies to protect confidential real data used for software development and testing	30%	23%
Q1c. My organization takes appropriate steps to protect the privacy rights of consumers, customers, employees and other individuals whose data is used for software development and testing.	31%	24%
Q1d. My organization takes appropriate steps to comply with privacy and data protection requirements.	44%	34%
Q1e. My organization has ample resources to ensure privacy and data security requirements are met in the application development and testing area.	30%	23%
Average	36%	27%

Q2. What types of real data do you use for development and testing purposes? Please check all that apply:	UK	US
Taxpayer or citizen records	15%	13%
Employee records	40%	45%
Customer records	61%	57%
Consumer lists	14%	15%
Credit cards	33%	42%
Other payment transactions	26%	35%
Vendor records	21%	30%
Other business confidential information	23%	45%
Average	29%	35%

Q3. In comparison to your organization's safeguarding of sensitive or confidential data in the production environment, which statement best describes your protection of real data used for development and testing?	UK	US
My organization uses the same safeguards when protecting sensitive or confidential data in both production and development.	52%	44%
My organization uses more stringent safeguards when protecting sensitive or confidential data in production than in development.	7%	8%
My organization uses less stringent safeguards when protecting sensitive or confidential data in production than in development.	41%	48%
Total	100%	100%

Q4a. How many data breaches has your organization experienced in the past 12 months?	UK	US
None [go to Q5]	25%	18%
1	50%	31%
2 to 5	22%	43%
More than 5	3%	8%
Total	100%	100%

Q4b. What were the reasons for the data breach? Please check more than one if you had two or more breaches in the past 12 months.	UK	US
Negligent insiders	57%	60%
Malicious insiders	31%	35%
3 rd party/Outsourcer	30%	41%
External attack by hackers	7%	7%
External attack by malware or bot.net	5%	10%
Social engineering, pretexting, or phishing	5%	5%
Other, please specify	1%	1%
Don't know	9%	27%
Total	145%	186%

Q5. How often do you require data for the application testing process?	UK	US
Daily	12%	20%
Weekly	42%	53%
Monthly	34%	18%
Quarterly	10%	9%
Other (please specify)	2%	0%
Total	100%	100%

Q6. How often would you say your organization experiences a business-changing event - e.g. this could be a new acquisition, a new product line or updated product pricing?	UK	US
Daily	1%	2%
Weekly	3%	5%
Monthly	15%	24%
Every other month	13%	18%
Every 3-6 months	13%	20%
Every 6-9 months	8%	9%
Every 9-12 months	16%	6%
About every 12 months or longer	31%	16%
Total	100%	100%

Q7. What data security measures do you have in place when testing?	UK	US
None	16%	20%
Internal firewalls	81%	78%
Training for those handling or using data	29%	17%
Access controls	30%	34%
Encryption technology	26%	31%
Data masking	28%	32%
Other (please specify)	2%	4%
Total	212%	215%

Q8a. What do you estimate is the size of the data sources typically used in the application testing process?	UK	US
Less than 1 terabyte	29%	21%
1-10 terabytes	44%	41%
10-50 terabytes	21%	34%
More than 50 terabytes	6%	4%
Total	100%	100%

Q8b. How often is that data refreshed?	UK	US
Every time	13%	20%
Frequently	50%	41%
Occasionally	29%	34%
Rarely	8%	4%
Total	100%	100%

Q8c. Is that data masked before it is used in testing?	UK	US
All production data is masked	10%	10%
A majority of production data is masked	8%	11%
A minority of production data is masked	10%	13%
No	72%	67%
Total	100%	100%

Q8d. Is that data reduced (subset)?	UK	US
Yes	30%	29%
No	70%	71%
Total	100%	100%

Q9. Do you send test data to a third party/outsourcer?	UK	US
No	40%	25%
Yes – only in-country	19%	17%
Yes – overseas	41%	58%
Total	100%	100%

Q10. Is the testing phase a bottleneck in your application development process?	UK	US
Yes	55%	57%
No	45%	43%
Total	100%	100%

Q11. Who is responsible for protecting data during the development and testing process? Please check the <u>one</u> responsible party within your organization.	UK	US
Business units sponsoring the development	31%	15%
Software development group	29%	7%
IT operations group	8%	18%
IT quality assurance group	5%	5%
IT security department	12%	21%
Privacy office	0%	1%
Corporate compliance and audit	0%	4%
No one party has responsibility	11%	20%
I don't know	3%	9%
Other (please specify)	1%	0%
Total	100%	100%

Q12. With respect to IT compliance needs, how important is the protection of real data used for development and/or testing purposes as compared to other compliance requirements?	UK	US
Very important	25%	27%
Important	35%	34%
Somewhat important	29%	22%
Not important	9%	13%
Irrelevant	2%	3%
Total	100%	100%

Demographics and Organizational Characteristics

D1 [UK]. Your current title (approximation)	UK	
Software engineer	23%	
Application developer	19%	
Quality assurance	15%	
Testing and development	12%	
Programmer	10%	
All other titles	21%	
Total	100%	

D1 [US]. Your current title (approximation)		US
Application developer		25%
Programmer		20%
Testing and development		15%
Quality assurance		11%
Software engineer		11%
All other titles		17%
Total		100%

D2. What organizational level best describes your current position?	UK	US
Senior Executive	0%	0%
Vice President	0%	1%
Director	8%	12%
Manager	23%	23%
Supervisor	21%	18%
Technical	35%	32%
Associate or staff	3%	5%
Administrative	0%	1%
Contractor	9%	8%
Other	1%	0%
Total	100%	100%

D3. Check the Primary Function where you reside within your organization.	UK	US
IT Operations	17%	13%
Application development	71%	75%
Compliance	6%	1%
Research	0%	0%
Human resources	0%	0%
Procurement	0%	2%
Security	5%	6%
Risk management	1%	2%
Privacy office	0%	0%
Total	100%	100%

D4. What industry best describes your organization's industry focus?	UK	US
Airlines	2%	2%
Automotive	0%	1%
Agriculture	0%	0%
Brokerage	2%	3%
Cable	0%	0%
Chemicals	1%	0%
Credit Cards	4%	4%
Defense	2%	2%
Education	5%	4%
Energy	1%	2%
Entertainment and Media	3%	3%
Central Government	9%	10%
Food Services	0%	0%
Healthcare	6%	7%
Hospitality & Leisure	3%	3%
Manufacturing	4%	4%
Insurance	2%	2%
Internet & ISPs	1%	1%
Local Government	6%	3%
Pharmaceuticals	3%	3%
Professional Services	3%	3%
Research	2%	2%
Retailing	7%	3%
Retail Banking	12%	14%
Telecommunications	4%	4%
Technology & Software	14%	18%
Transportation	2%	2%
Wireless	1%	1%
Total	100%	100%

D5 [UK]. Is your company publicly traded?	UK	
Yes, FTSE 100	21%	
Yes, FTSE 500	32%	
Yes, not ranked in the FTSE	15%	
No	32%	
Total	100%	

D5 [US]. Is your company publicly traded?		US
Yes, Fortune 100		12%
Yes, Fortune 500		24%
Yes, not ranked by Fortune		21%
No		42%
Total		100%

D6 [UK]. What is the approximate total revenue of your organization in the past 12 months?	UK	
Below £10 million	6%	
£10 to £100 million	9%	
£101 million to £1 billion	13%	
£1 to £10 billion	21%	
£11 to £20 billion	32%	
More than \$20 billion	19%	
Total	100%	

D6 [US]. What is the approximate total revenue of your organization in the past 12 months?		US
Below \$10 million		3%
\$10 to \$100 million		8%
\$101 million to \$1 billion		14%
\$1 to \$10 billion		18%
\$11 to \$20 billion		30%
More than \$20 billion		28%
Total		100%

D7. What is the worldwide headcount of your organization?	UK	US
Less than 500 people	7%	4%
500 to 1,000 people	8%	9%
1,001 to 5,000 people	27%	20%
5,001 to 25,000 people	30%	23%
25,001 to 75,000 people	21%	30%
More than 75,000 people	7%	14%
Total	100%	100%

All responses are completely confidential.
Please contact research@ponemon.org or call us at 800.877.3118 if you have any questions.

Ponemon Institute

Advancing Responsible Information Management

Ponemon Institute is dedicated to independent research and education that advances responsible information and privacy management practices within business and government. Our mission is to conduct high quality, empirical studies on critical issues affecting the management and security of sensitive information about people and organizations.

As a member of the **Council of American Survey Research Organizations (CASRO)**, we uphold strict data confidentiality, privacy and ethical research standards. We do not collect any personally identifiable information from individuals (or company identifiable information in our business research). Furthermore, we have strict quality standards to ensure that subjects are not asked extraneous, irrelevant or improper questions.