



DATA ARCHIVING

A window on the life of a company

Phil Manchester finds out why good management of information is important

Data are at the heart of all regulatory compliance. The way information is generated, stored and processed defines the activities of every organisation and provides the evidence of good, or bad, governance. It is no surprise, therefore, that all the regulatory initiatives currently affecting businesses around the world place great emphasis on how data is managed.

High quality data are required for reporting under a wide range of regulations – from the general International Financial Reporting Standard (IFRS) and US Sarbanes-Oxley Act to specific financial sector regulations such as the Basel II capital accord and the European Union's market in financial instruments directive (Mifid).

Clearly, transactional data are a priority because they are a direct representation of business activity. But increasingly, regulatory authorities want to see other kinds of less formal data. Text data, e-mail – and even voicemail – are on the list of items to be archived under some regulatory regimes.

The scale of data archiving is enormous. IDC, the research company, estimates businesses around the world generated 161 exabytes (an exabyte is a billion gigabytes) of data in 2006*.

Not all data come under a compliance regime. But a significant proportion will.

"Merrill Lynch generates 85m instant messages every quarter. This means for compliance purposes it must store around 1bn messages over a period of three years," says Bill Lyons, chief execu-

tive officer of AXS-One, a US-based data archiving specialist. He says the regulators' growing demands for a rapid response means organisations must have ways to get at archived data quickly.

"A typical large financial institution will carry out 800 searches a year in support of compliance and litigation and we expect this to triple in the next few years," says Mr Lyons. "If you don't have effective archiving it can be very costly to respond to the demand for searches. The way we approach archiving is to aim to achieve Google-like speeds on data archive searching."

Other data management software vendors have responded to the demand and built features into their software to handle the requirements of compliance. Leading enterprise software vendors SAP and Oracle, for example, include compliance features in their products.

Organisations must also deal with the problem of data quality – especially when they need to draw data together from different sources.

Informatica, a data integration specialist, says companies must ensure they have a clear infrastructure for data integration to ensure data quality.

"The chief financial officer wants to know that when he signs off reports for regulatory authorities, he is signing off high-quality data," says Tommy Drummond, vice-president of product marketing for Informatica's data quality services.

"Companies need to clean up consolidated data to meet their compliance obligations so they can make sure the



Off the shelf: regulators' demands for rapid responses mean groups must get to archived data quickly Getty

regulator has confidence in the data," he says.

M&S Money, an Informatica customer and a wholly-owned subsidiary of HSBC bank, was faced with a significant data "cleansing" exercise when it was obliged to meet the Basel II regulations on risk management. But it also saw the improvements in data quality in a broader context – beyond simply meeting the rules.

"The main reason for our data quality drive was for Basel II," says Neil Hershaw, information management officer at M&S Money. "We have to prove to the Financial Services Authority (FSA) that we have the right processes in place and our data are up to scratch. Although this was the main driver, we wanted to get the most out of the exercise and gain from having high data quality."

The bank set up its information management office in 2005 to see how it could improve data quality ahead of the Basel II deadline. Although it had used an Oracle data warehouse for several years, there were concerns about data quality.

"This approach to data quality was new to us and we did not have the processes in place to ensure it," says Mr Hershaw. "The problem in any company is that everyone assumes the data are OK because they only see them as a summation. It is not until you get down to the low level that you can see where the faults are. And you need software to plough through and carry out the checks."

Mr Hershaw says there is a range of checks that need to be applied to ensure data quality. These include general completeness and con-

sistency checks in addition to more detailed checks, such as making sure a post-code matches an address or a birth date is in range.

"The Informatica data quality checking software we use enables us to code up our rules and check that the data conform. The software produces a scorecard to show us how accurate the data are and we can use this as a basis for making corrections," he says.

Mr Lyons of AXS-One says organisations go through three phases in building high-quality data archives: "Typically they begin with the aim of improving operational efficiency. Then they move on to meet compliance rules. Finally they ask how they can reuse the data for management information."

*www.emc.com/about/destination/digital_universe/.