Protecting Data Privacy using IBM InfoSphere Software (Optim & Guardium)

Stephen Tallant

September 2010

InfoSphere: Collaborative Information Governance

Reusability and consistency

- Shared metadata and policies

Breadth of portfolio

- Three core information governance disciplines

Modular deployment entry points

- Supports business and IT priorities

Flexible support for enterprise environments

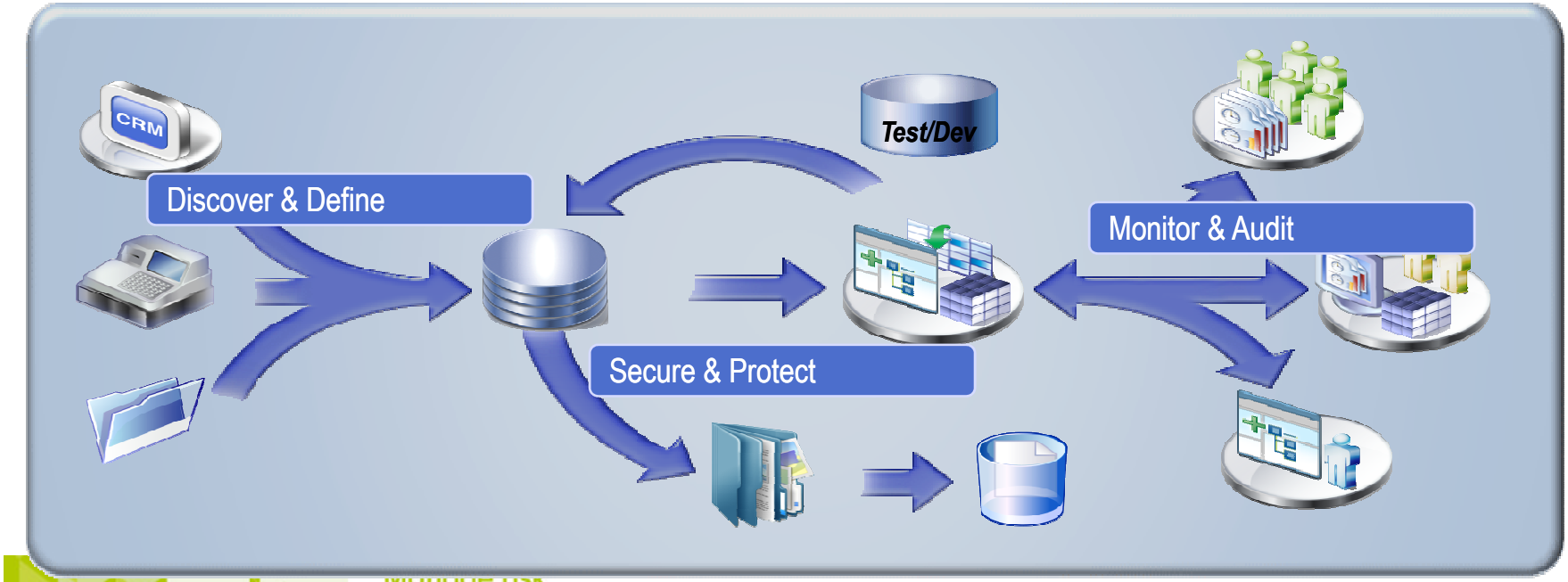
- Open technology for heterogeneous support

Single solution provider to Optimize the Information Supply Chain

Manage risk
in a complex world

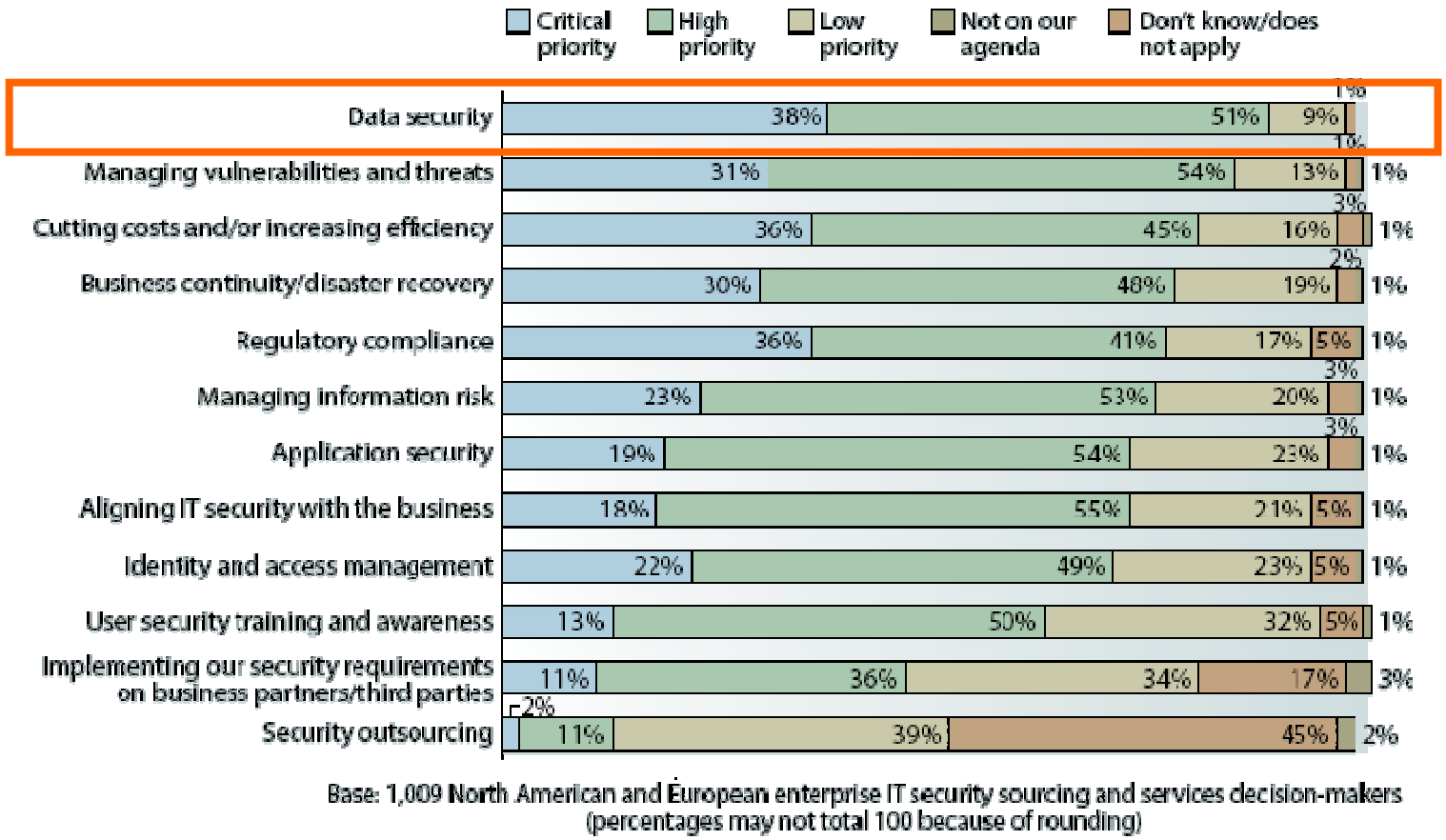
Securing and Protecting Your Information Supply Chain

- Understanding the “what & where” of enterprise data
- Protecting the data across the enterprise, both internal and external threats
- Knowing who’s accessing your data when, how and why
- Monitoring and reporting on database access for audit purposes



Survey: Protecting Data Remains Important

"Which of the following initiatives are likely to be your organization's top IT security priorities over the next 12 months?"



Source: Forrester Research, Inc. Jonathan Penn, "The State Of Enterprise IT Security And Emerging Trends: 2009 To 2010" – January 2010

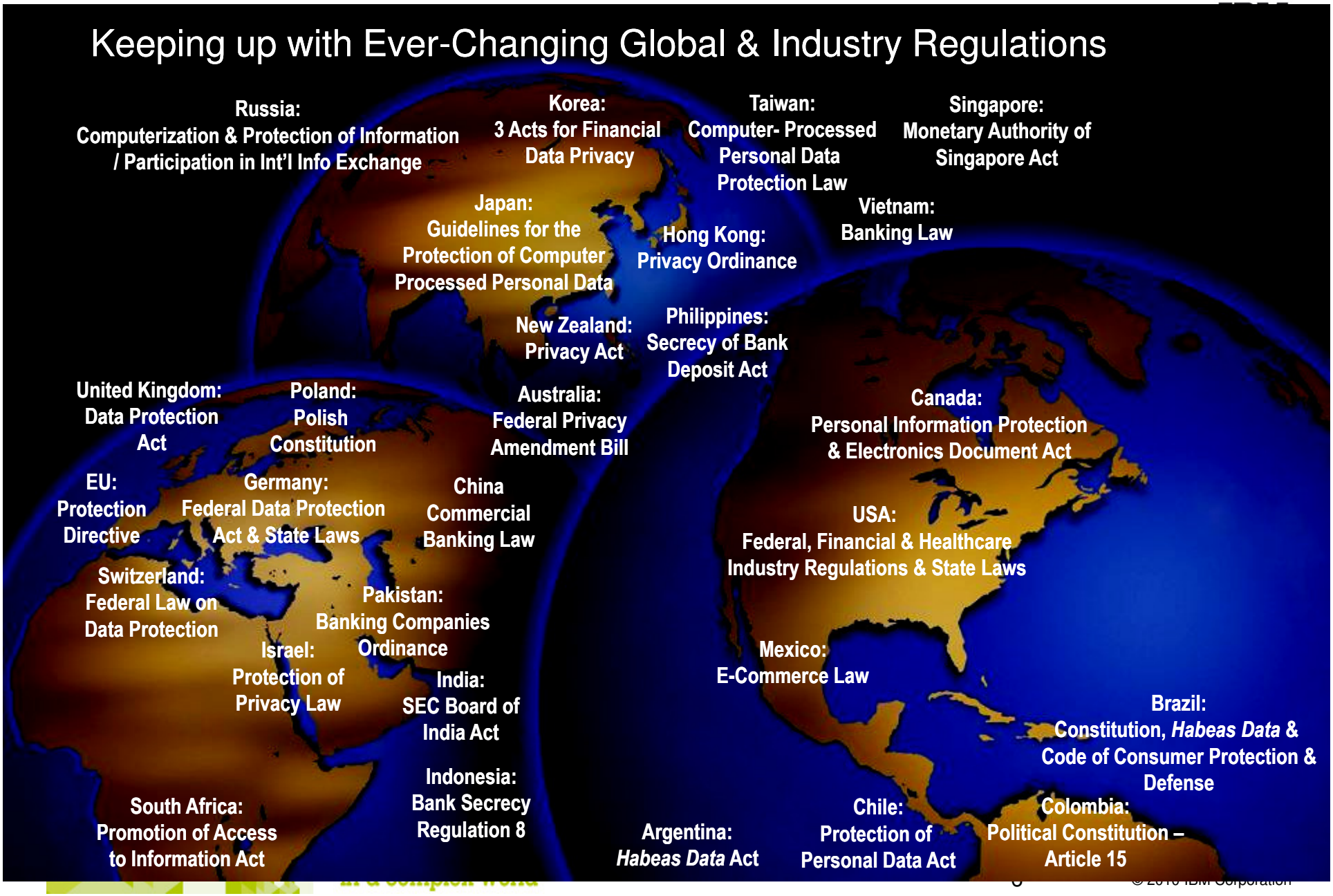
Organizations have multiple data protection challenges

- *Limited time, lots of regulation, growing costs of compliance*
  - Organizations under time pressure to show compliance progress to the business
  - Meeting privacy regulatory requirements in a timely and cost-effective manner

- *Requirements for privacy/security by user role add complexity*
  - Ensuring access to enterprise data adheres to the various job roles (Billing clerk vs. Doctor) for sensitive data fields
  - Ad-hoc solutions often replicate sub-sets of information to meet role requirements

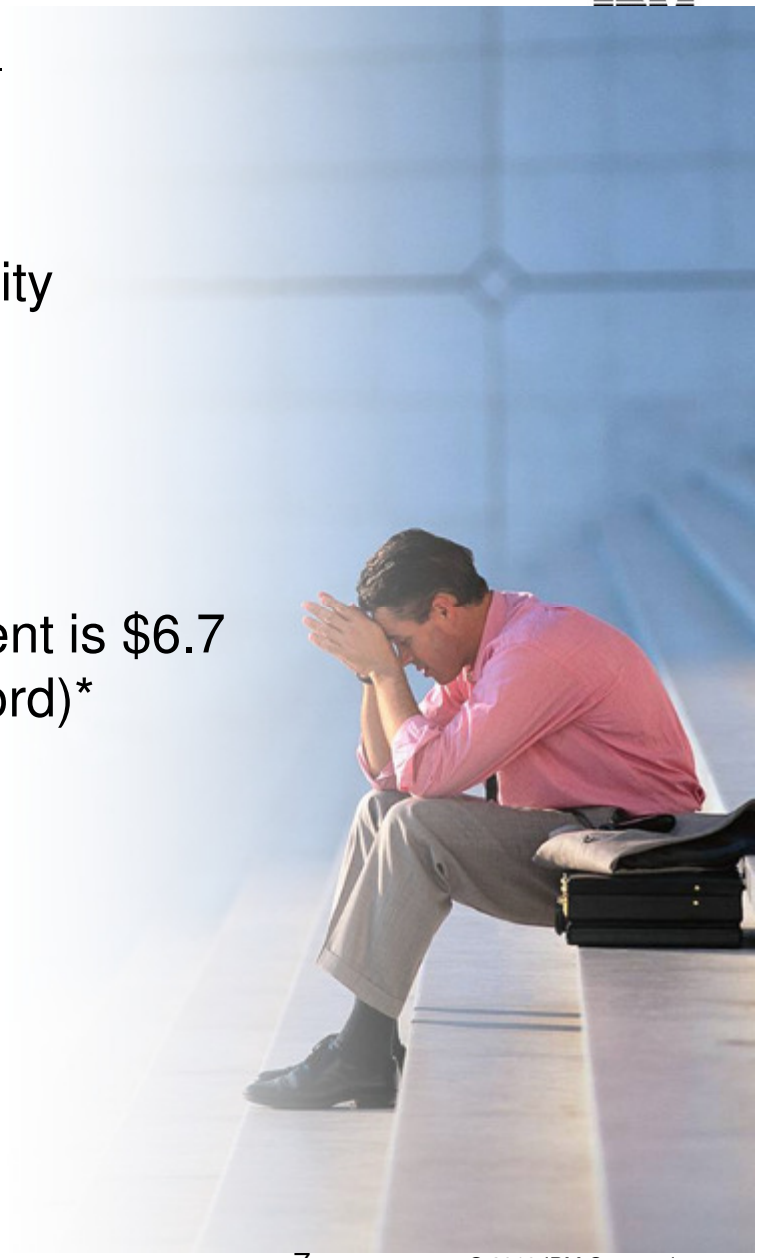
- *Manual approaches lead to higher risk and inefficiency*
  - Ineffective home-grown solutions applied to mask structured and unstructured data
  - Complex, manual processes used to identify sensitive data, perform security audits and track user access

Keeping up with Ever-Changing Global & Industry Regulations



What's at Stake?

- Damage to company reputation
  - "Brand equity" damage; negative publicity
  - Loss customer loyalty
- A privacy breach – or the threat of one
  - Intellectual property loss
  - Increased operations cost
    - Average cost of a data breach incident is $6.7 million ($204 per compromised record)*
- Loss of revenue & share price erosion
- Audits and the possibility of being fined



* Sources: Ponemon Institute, 2009

What's the Risk?

Confidential data inadvertently exposed or otherwise available to unauthorized viewers.

February 2010: *About 600,000 customers of a major NYC bank received their annual tax documents with their Social Security numbers (combined with other numbers & letters) printed on the outside of the envelope.*

SQL injection is fast becoming one of the biggest & most high profile web security threats.

July 2010*: Hackers obtained access to the user database and administration panel of a popular website by exploiting several SQL injection vulnerabilities. The exposed data included user names, passwords, e-mail addresses and IPs.*

Unprotected test data sent to and used by test/development teams as well as third-party consultants.

February 2009: *An FAA server used for application development & testing was breached, exposing the personally identifiable information of 45,000+ employees.*

Confidential data that should be redacted can be hidden or embedded

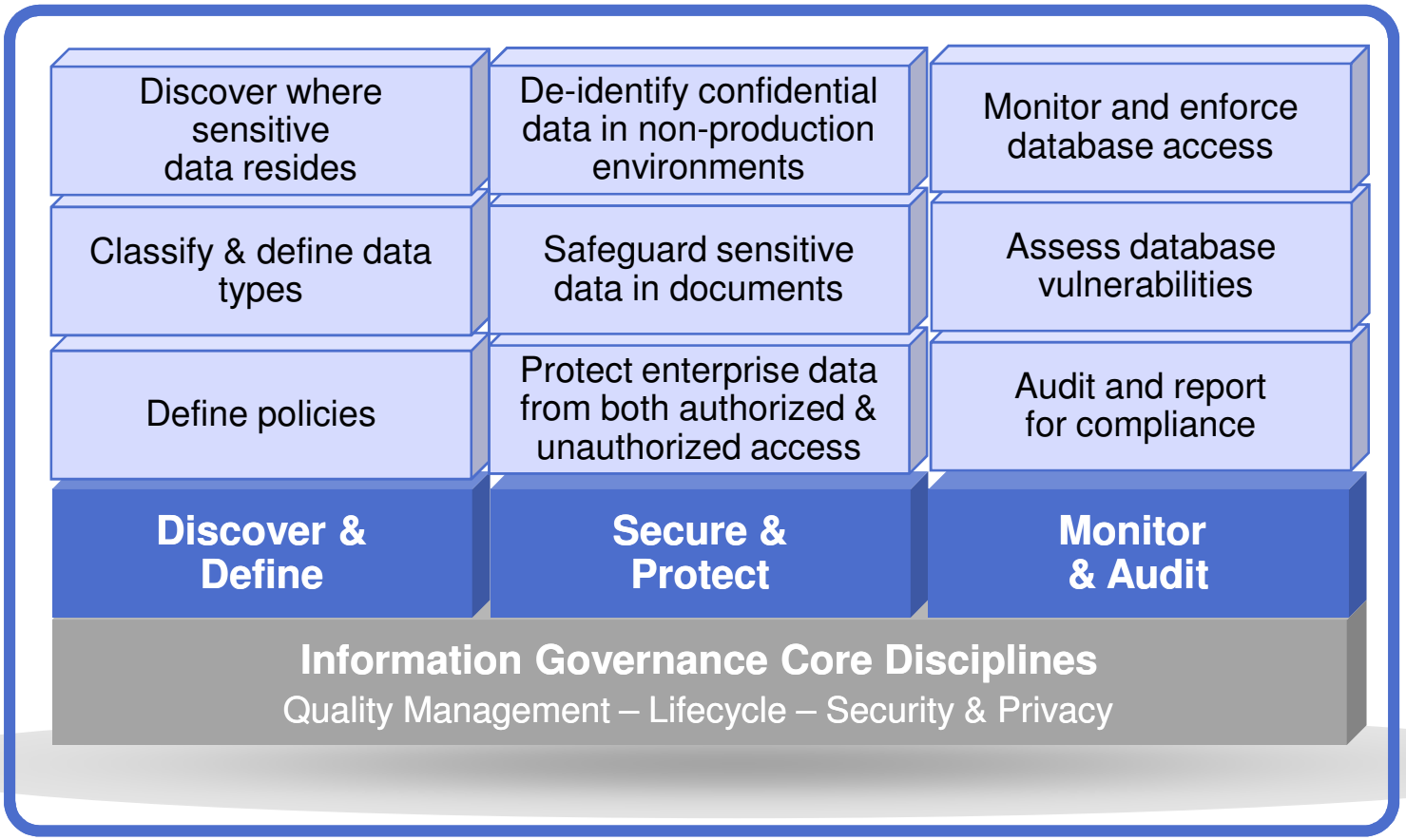
April 2010: *A PDF of a subpoena in the case of "United States vs. Rob Blagojevich" was posted to public website. However, the "redacted" text simply had black box placed on top to hide the content – the actual text was still available.*

Can Today's Organizations Successfully Protect Their Information?

- Where does your sensitive data reside across the enterprise?
- How can your data be protected from both authorized and unauthorized access?
- Can your confidential data in documents be safeguarded while still enabling the necessary business data to be shared?
- How can access to your enterprise databases be protected, monitored and audited?
- Can data in your non-production environments be protected, yet still be usable for training, application development and testing?

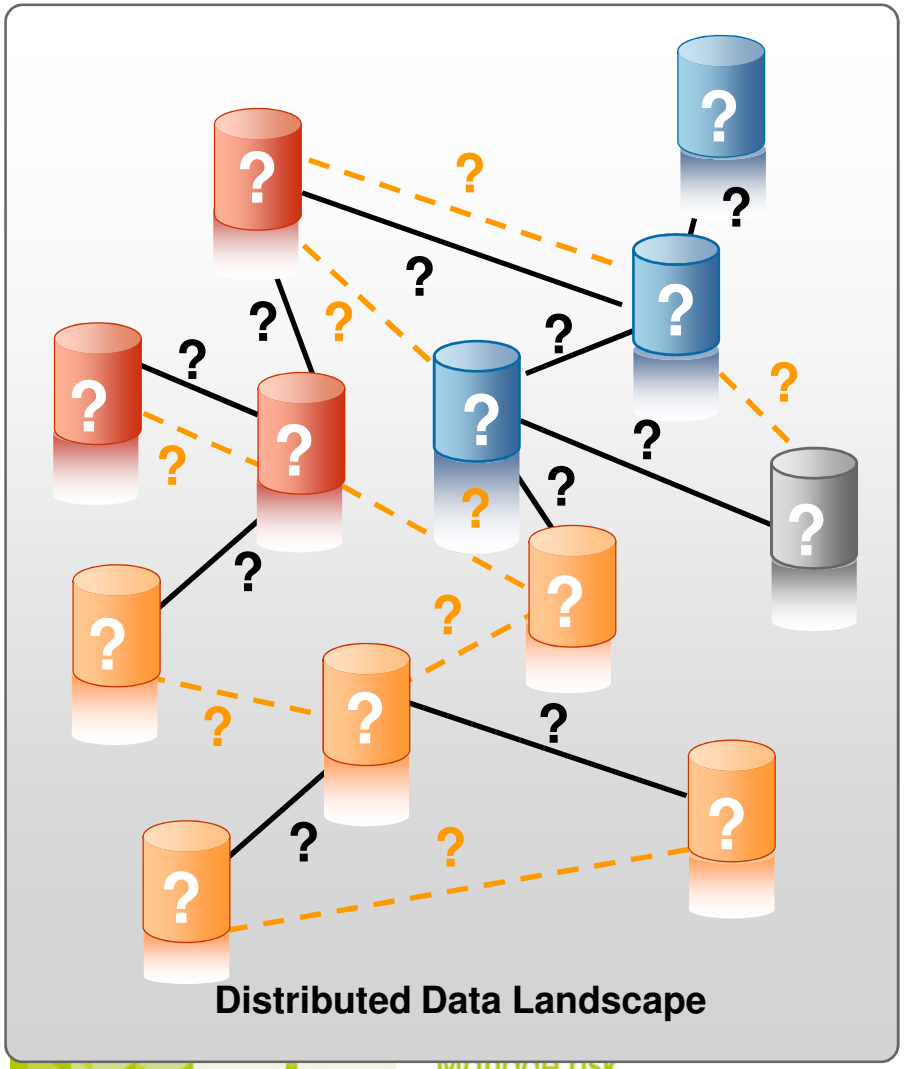
> *"Larry Ponemon, founder of the group that bears his name, said that survey shows a shift in the way C-level executives think about security software. Investing in data protection, he said, is now seen as less expensive than recovering from a data breach.* **-- InformationWeek**

Protecting Information Security & Privacy Across the Enterprise



You Can't Govern what You Don't Understand

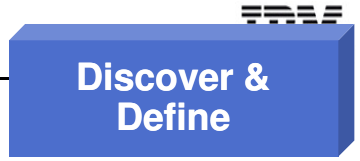
Discover & Define



Distributed Data Landscape

- Data can be distributed over multiple applications, databases and platforms
  - Where are those databases located?
- Complex, poorly documented data relationships
  - Which data is sensitive, and which can be shared?
  - Whole and partial sensitive data elements can be found in hundreds of tables and fields
- Data relationships not understood because:
  - Corporate memory is poor
  - Documentation is poor or nonexistent
  - Logical relationships (enforced through application logic or business rules) are hidden

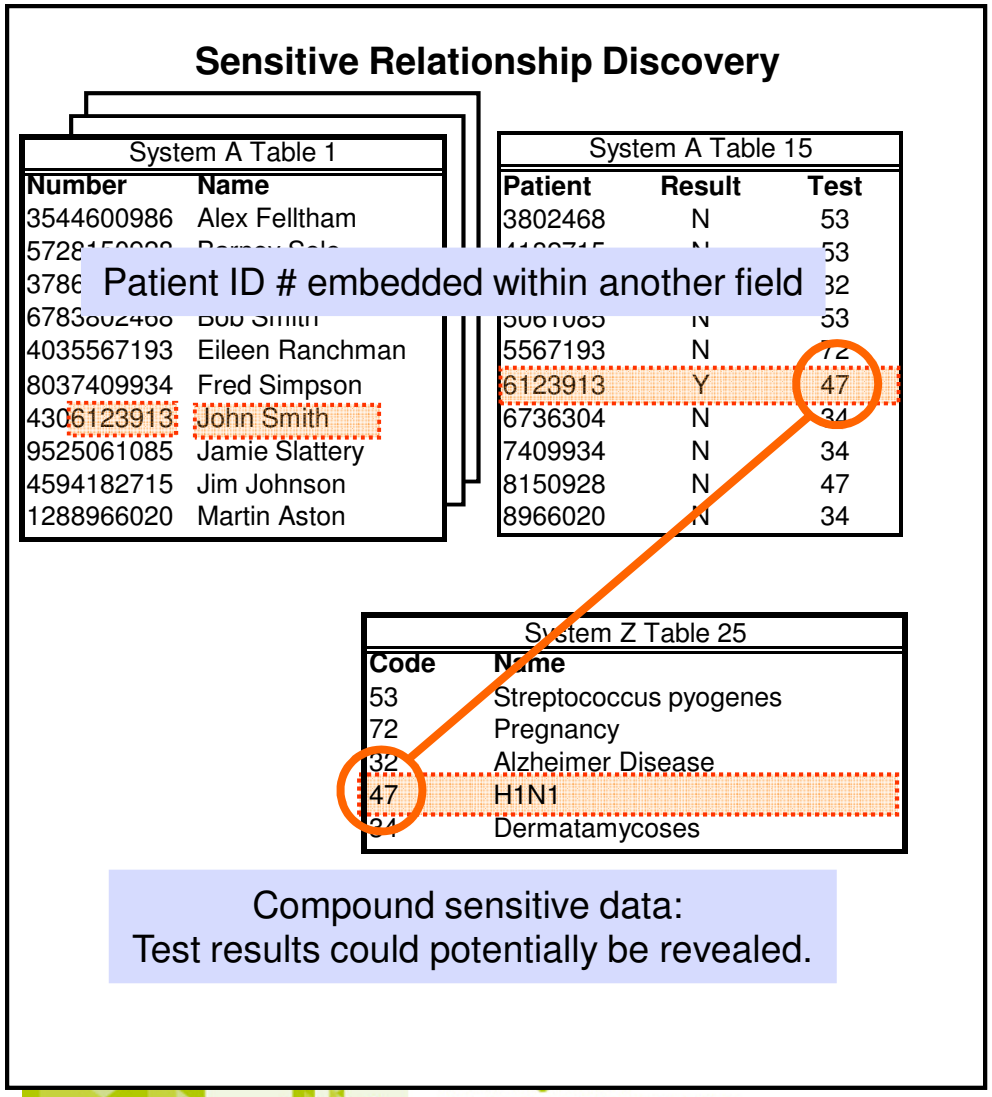
Locate Data and Data Relationships



- Locate and inventory the databases across the enterprise
- Identify sensitive data and classify
- Understand relationships required for identifying compound sensitive data
- Define and document the privacy & masking rules and propagate to ensure sensitive data will be protected
- Document and manage ongoing data masking requirements

Discover How Data is Related and Where Sensitive Data May Be Hidden

Discover & Define



- Relationships and sensitive data can't always be found just by a simple data scan
  - Sensitive data can be embedded within a field
  - Sensitive data could be revealed through relationships across fields & systems
- When dealing with hundreds of tables and millions of rows, this search is complex – you need the right solution

Protecting Data is Both an External and Internal Issue

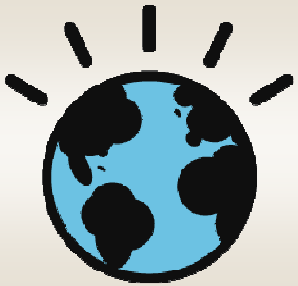
Secure & Protect

- Prevent “power users” from abusing their access to sensitive data (separation of duties)
  - DBA and power users
- Prevent authorized users from misusing sensitive data
  - For example, third-party or off-shore developers
- Prevent intrusion and theft of data
  - For example, someone walking off with a back-up tape
  - Hacker
  - Database vulnerabilities (user id with no password or default password)

Protection of data requires a 360-degree strategy

- Secure sensitive data values
  - Across both structured and unstructured
- De-identify data
  - Restricted data sharing with 3rd parties
  - Generation of fictionalized test data for non-production
  - Support off-shore deployment model
- Stop unauthorized data access
  - Render data useless via encryption
  - Lock down SQL to prevent SQL injection
  - Block suspicious network traffic

Security makes it possible for us to take risk, and innovate confidently.



Protect Sensitive Data Values within Documents

Secure & Protect

- Redact (or remove) sensitive unstructured data found in documents and forms, protecting confidential information while supporting the need to share critical business information
  - Support compliance with industry-specific and global data privacy requirements or mandates
- Leverage an automated redaction process for speed, accuracy and efficiency
  - Ensure hidden source data (or metadata) within documents is redacted as well
- Prevent unintentional disclosure by using role-based masking to confidently share data
- Ensure multiple file formats are support, including PDF, text, TIFF and Microsoft Word documents

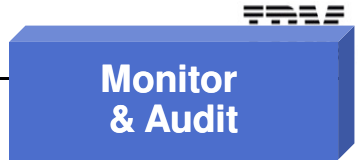
De-identify Data in Non-Production Environments without Impacting Test & Development

Secure & Protect

- Mask or de-identify sensitive data elements that could be used to identify an individual
- Ensure masked data is contextually appropriate to the data it replaced, so as not to impede testing
  - Data is realistic but fictional
  - Masked data is within permissible range of values
- Support referential integrity of the masked data elements to prevent errors in testing

Personal identifiable information is masked with realistic but fictional data for testing & development purposes.

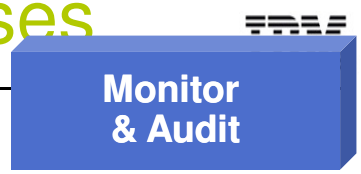
What happens with security complacency



- Not being able to report compliance can lead to regulatory fines
  - No audit report mechanism
  - No fine grain audit trail of database activities
- Don't know if there is a data breach until its too late
  - Lack of awareness of suspicious access patterns
  - On-going vs. single-invent: problems identifying patterns of unauthorized use
- Not able to monitor super user activity to ensure data security standards
  - Unable to detect intentional and unintentional events

> "Most organizations do not have mechanisms in place to prevent database administrators and other privileged database users from reading or tampering with sensitive information [in business applications]…Fewer than two out of five respondents said they could prevent such tampering by super users.
>
> **-- Independent Oracle User Group**

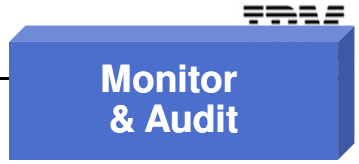
Streamline and simplify compliance processes



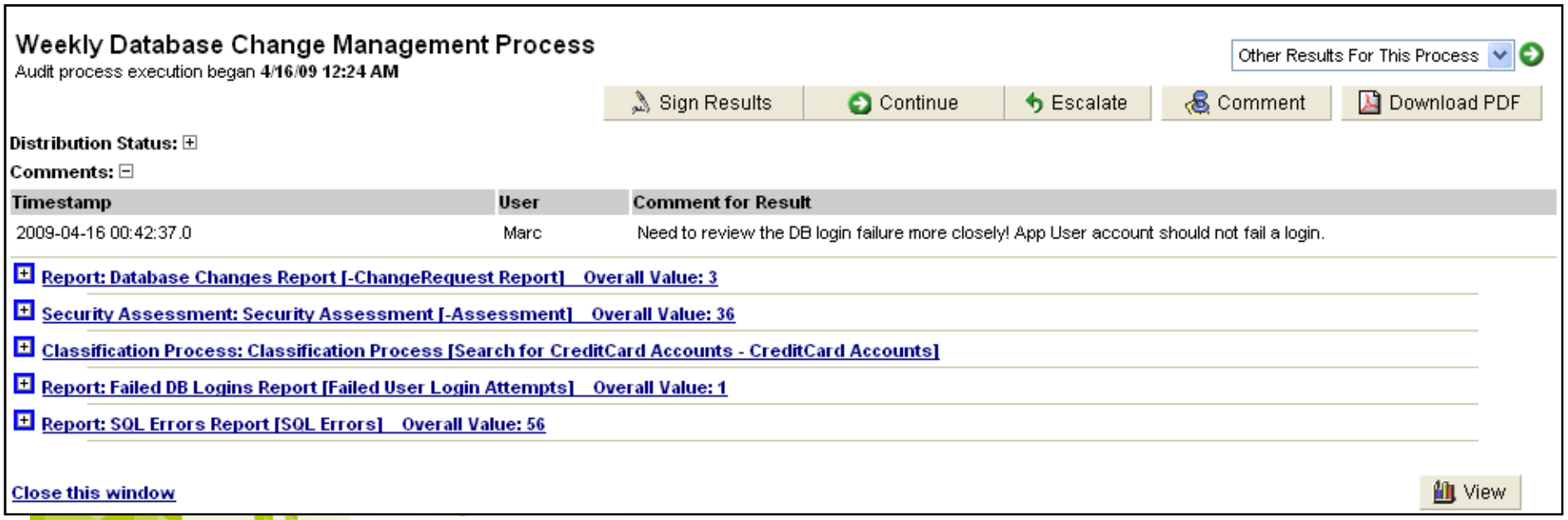
- Alerts of suspicious activity
- Audit reporting and sign-offs
  - user activity,
  - object creation
  - Database configuration
  - Entitlements
- Separation of duties – creation of policies vs. reporting on application of policies
- Trace users between applications, databases
- Fine grained-policies
- Sign-off and escalation procedures
- Integration with enterprise security systems (SIEM)

Manage risk
in a complex world

Simplify and Streamline Audit Process



- Generate audit reports and distribute to oversight team
  - Electronic sign-offs
  - Escalations, comments and exception handling
- Document oversight processes, addressing auditors' requirements
- Store audit process results with audit data in secure audit repository



IBM Provides the Expertise to Protect and Secure Data

- A comprehensive Information Governance strategy addresses the need for data protection, security and privacy to safeguard corporate information assets
- IBM's solutions enable our clients to create and maintain trusted information infrastructures, protect high-value enterprise databases and safeguard sensitive data throughout the enterprise
- Armed with trusted information, our clients can successfully transform their businesses to deliver new value, control cost, and mitigate compliance risks

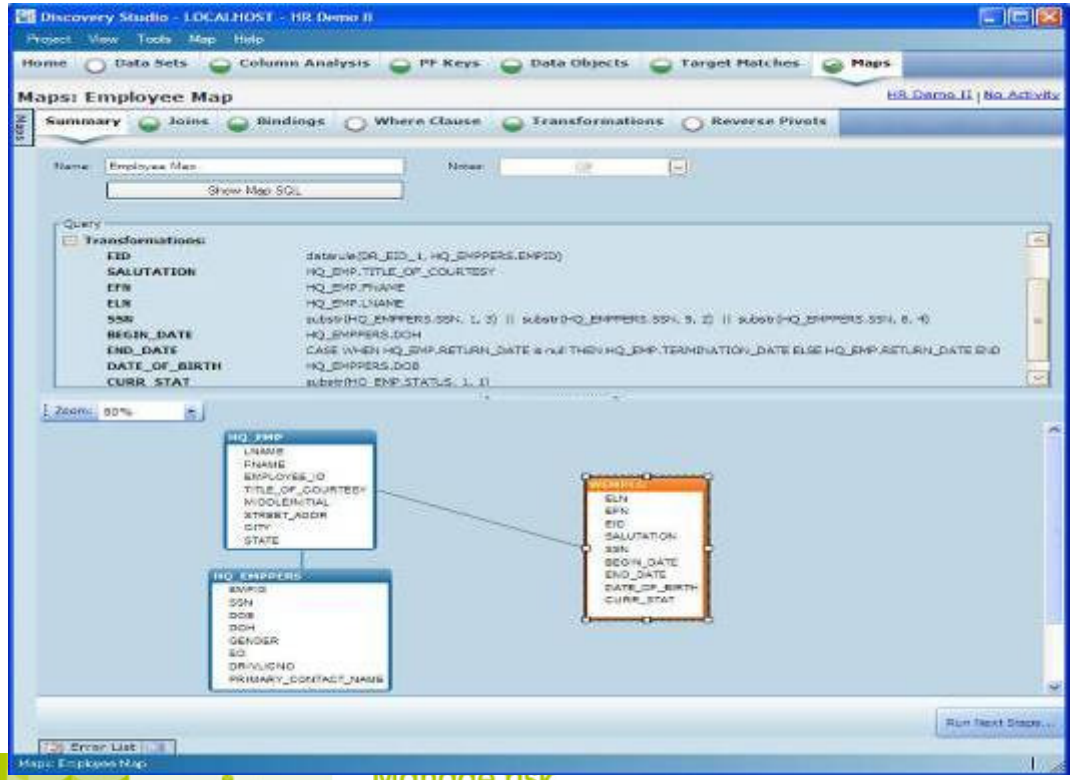
> "The top challenge for 43% of CFOs is improving governance, controls, and risk management
>
> *CFO Survey: Current state & future direction,* IBM Business Consulting Services

Manage risk
in a complex world

IBM InfoSphere Discovery

Discovery

Accelerate project deployment by automating discovery of your distributed data landscape



Requirements

- Define business objects for archival and test data applications
- Discover data transformation rules and heterogeneous relationships
- Identify hidden sensitive data for privacy

Benefits

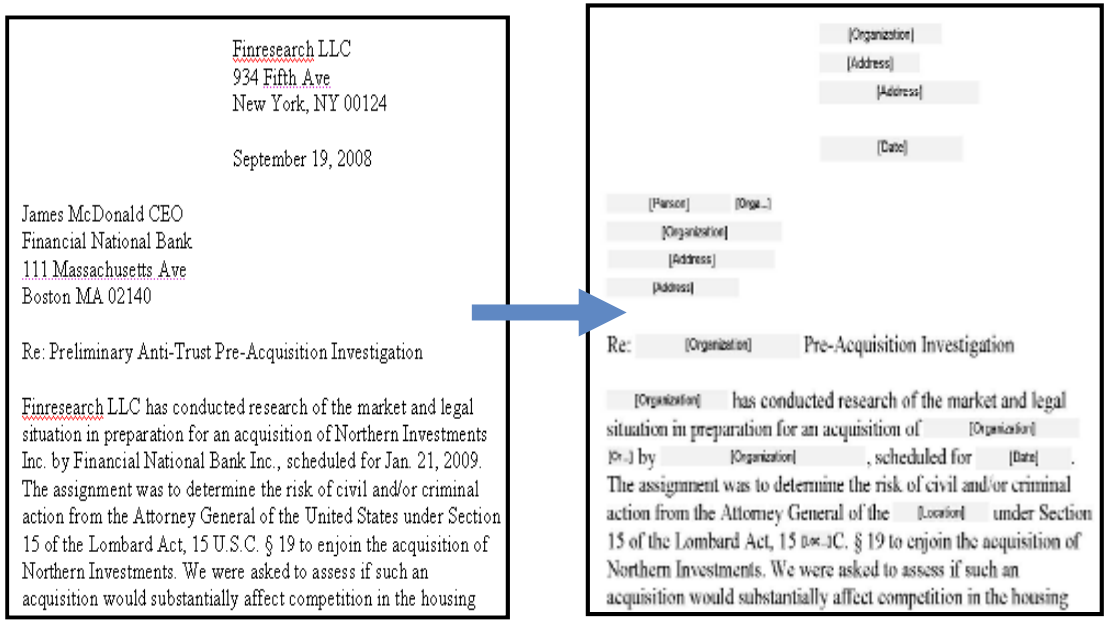
- Automation of manual activities accelerates time to value
- Business insight into data relationships reduces project risk
- Provides consistency across information agenda projects

Manage risk
in a complex world

IBM InfoSphere Guardium Data Redaction

Data Redaction

Protect sensitive unstructured data in documents and forms



Finresearch LLC
934 Fifth Ave
New York, NY 00124

September 19, 2008

James McDonald CEO
Financial National Bank
111 Massachusetts Ave
Boston MA 02140

Re: Preliminary Anti-Trust Pre-Acquisition Investigation

Finresearch LLC has conducted research of the market and legal situation in preparation for an acquisition of Northern Investments Inc. by Financial National Bank Inc., scheduled for Jan. 21, 2009. The assignment was to determine the risk of civil and/or criminal action from the Attorney General of the United States under Section 15 of the Lombard Act, 15 U.S.C. § 19 to enjoin the acquisition of Northern Investments. We were asked to assess if such an acquisition would substantially affect competition in the housing

Before

After

Requirements

- Protect unstructured data in textual, graphical and form based documents
- Control data views with user role policies
- Automate batch workflow process with optional human review

Benefits

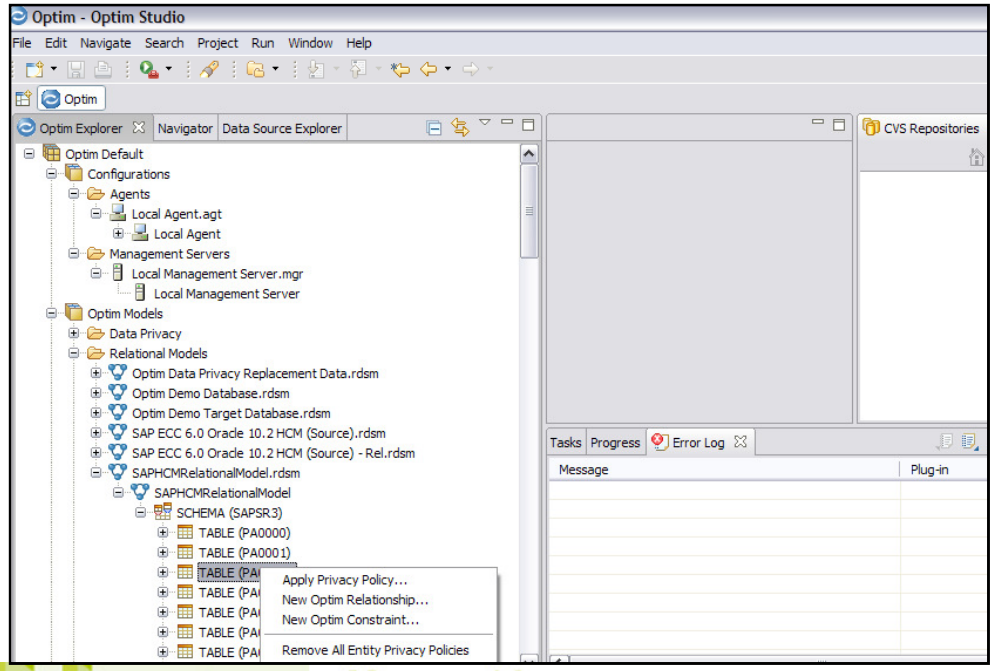
- Prevent unintentional data disclosure
- Comply with regulatory and corporate compliance standards
- Increase efficiency and reduce risk via automation

Manage risk
in a complex world

IBM InfoSphere Optim Data Privacy Solution

Data Privacy

De-identify sensitive information with realistic *but fictional* data for testing & development purposes



Requirements

- Protect confidential data used in test, training & development systems
- Implement proven data masking techniques
- Support compliance with privacy regulations
- Solution supports custom & packaged ERP applications

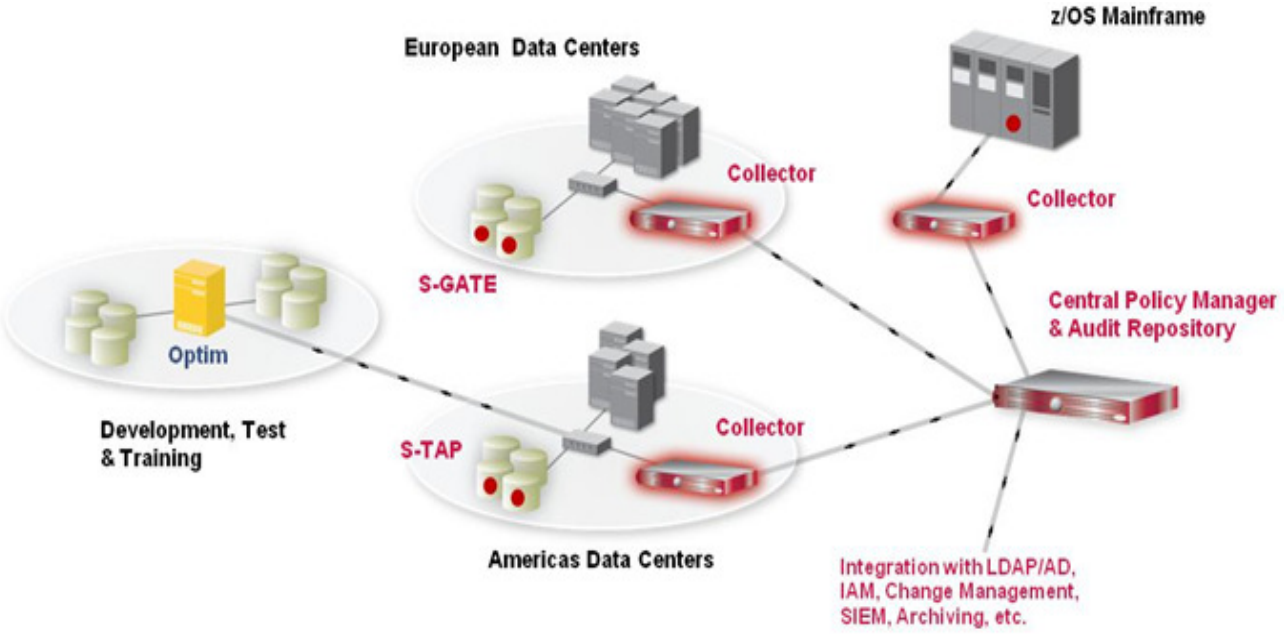
Benefits

- Protect sensitive information from misuse and fraud
- Prevent data breaches and associated fines
- Achieve better data governance

IBM InfoSphere Guardium

Guardium

Database Protection and Compliance Made Simple



Requirements

- Continuous, real-time database access and activity monitoring
- Policy-based controls to detect unauthorized or suspicious activity
- Vulnerability assessment, change auditing and blocking

Benefits

- Assure compliance with regulatory mandates
- Protect against threats from legitimate users and potential hackers
- Minimize operational costs through automated and centralized controls

Manage risk
in a complex world

Success: Leading Global Household Goods Manufacturer Protects the Privacy of HR data within Non-Production systems

Challenge

- This leading household goods manufacturer needed to **consolidate multiple worldwide instances** of the SAP Human Capital Management application.
- As they created their testing environment, the client wanted to “de-identify” their SAP HCM data so that developers were not using **confidential employee HR data in their test environments**.

Solution

- **IBM InfoSphere Optim Data Privacy Solution for SAP Applications**

Business Benefits

- **Reduced time** to manually code the data scrambling routines.
- **Implemented data masking solution**, as part of overall support data governance strategy
- **Protected confidential employee information** within the testing and development environments, ensuring privacy of HR and payroll information
- Deployed data masking solution **quickly and efficiently**, using both out-of-box definitions as well as custom de-identification routines

Manage risk
in a complex world

Success: Large Insurance Organization Meets PCI DSS Compliance Requirements

Challenges

- **Meet compliance requirements for PCI DSS** (Payment Card Industry Data Security Standard) for content management of historical documents and forms
- **Diverse groups need access to different information** in documents which contain personal health information (PHI) and confidential financial information (credit card numbers)
- **Replace current** cumbersome, lengthy **manual process** to redact forms and documents and minimize risk.

Solution

- **IBM InfoSphere Optim Data Redaction**

Business Benefits

- Boost time-to-value with quick implementation and high accuracy rates for redaction candidates.
  - **97% accuracy**
- **Satisfy compliance requirements** in a timely manner
- **Increase efficiency and minimize risk of omissions** with automated identification and redaction of sensitive data

"We are thoroughly impressed with IBM Optim Data Redaction, its capabilities and accuracy rates. This technology is helping us comply with PCIDSS (Payment Card Industry Data Security Standard) requirements for historical content management of documents and forms."

Manage risk
in a complex world

Success: Leading Technology Company Simplifies Enterprise Security

Challenges

- Improve database security for SOX, PCI & SAS70
  - Environment: Oracle & SQL Server on Windows, Linux; Oracle E-Business, JD Edwards, Hyperion plus in-house applications
- Simplify & automate compliance controls
  - Previous solution consisted of traces & auditing with in-house scripts, which impacted DBA resources, and lead to massive data volumes, supportability issues and SOD issues

Solution

- **IBM InfoSphere Guardium**

Business Benefits

- **Enterprise-class scalability**, deployed to 300 DB servers in 10 data centers in 12 weeks (deployed to additional 725 database servers in phase 2).
- Addressed critical needs for automated compliance reporting; real-time alerting; and centralized cross-DBMS policies.
- Closed-loop change control with Remedy integration

"The Guardium architecture offers a noninvasive, network-based, database-independent platform for continuously monitoring and analyzing database traffic in real time to help immediately identify unauthorized or suspicious activities."