Bridging Strategy and Data

Practical Data Masking:

How to address Development and QA teams' 7 most common data masking related reactions and concerns

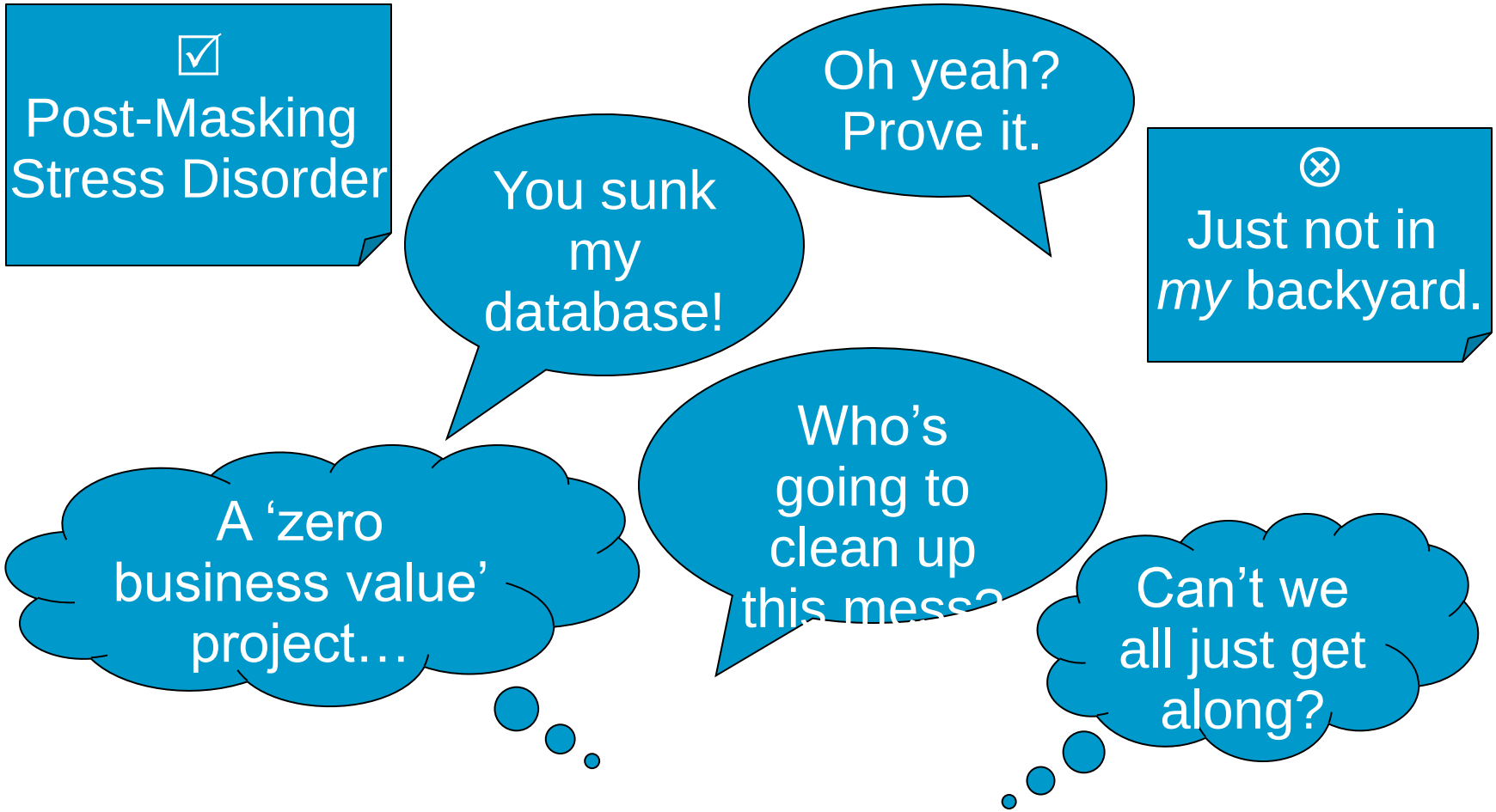
Presented to: New York Oracle User's Group
December 8th, 2010

Data Security Risk

	Description
Public	A data element that is clearly identified as sensitive data because, all on its own, it refers to information which can be used to distinguish an individual's identity. **Examples:** Name, Phone number, Street Address, Zip+4
Internally Identifying	Information is identifying to a company employee **Examples:** Extension, Company Id Number
Inference	A data element that, when used in conjunction with other data elements reveals the identity of an individual (can also be applied to business secrets.) **Example:** Last Name and SSN
Derived	A data element that contains a unique fact **Examples:** An application containing HR information may have one individual with a $1 salary. This could identify the CEO who may have reduced his salary as part of a cost-savings. Or a holding with a very large stock price.
Exploitable	A data element not identifying but can be directly exploited **Examples:** Credit Card Numbers, Checking Account Numbers, SSN

Practical Data Masking

How to address Development and QA teams' 7 most common data masking related reactions and concerns:



☑ Post-Masking Stress Disorder

1. Helping your loved ones cope with the loss of their real production data

Challenge

- *"We won't be able to test! The application won't work. I can't do my job like that..."*

Solution

- Make target data look and act realistic.
- *"Let us show you – just give us a sandbox, then check out the results before giving approval."*

Benefit

- Empower development teams in the process.
- Provide development teams with usable data.

MASK INTELLIGENTLY

☑ Post-Masking Stress Disorder

1. Helping your loved ones cope with the loss of their real production data

EXECUTE DATA MASKING PILOT	START	END	OWNER
Prep for Masking Pilot			
Select pilot applications and environment			
Deliver On-boarding packet to pilot Application Managers			
Select target data, test scripts and cases			
Set up sandbox environment for pilot			
Execute Smoke Test			
Back-up environment			
Run jobs (unmasked)			
Mask data and run jobs			
Compare results			
Restore environment			
Refine Masking Rules			
Develop masking jobs for new Sensitive Data elements			
Repeat Smoke Test if desired			
Execute Integration Test			
Back-up environment			
Mask data and run jobs			
Validate results, correct any issues			
Restore environment			
Execute Final Masking			
Mask data			
Transition			
Resolve any outstanding issues from Final Masking			
Invoke process for sustaining masked environment			
Support turnover process for new Sensitive Data data			

MASK INTELLIGENTLY

2. When it comes to meeting application testing requirements, referential integrity is just the tip of the iceberg.

Challenge

- *"These applications need to talk to each other even after they're masked."*

Solution

- Determine which systems:
  - Must be masked in synch.
  - Need to be masked first, then used to feed downstream applications.
- Identify data elements that must be preserved.
- Address interrelated fields.
- Select the best technique to mask each data element (context-dependent.)

Benefit

- Applications work and interact seamlessly.

2. When it comes to meeting application testing requirements, referential integrity is just the tip of the iceberg.

Framework	Definition	Features	Examples
Secure Lookup	• A proprietary algorithm, repeatable but unbreakable. • Used to assign a realistic value from a pre-defined lookup values. • While DMsuite comes with pre-configured secure lookups, a user can generate their own secure lookup values and easily import them into DMsuite.	• Irreversible • Duplicates Intended	•Names •Addresses
Segmented Mapping	• Create masked values by dividing target value into separate segments and masking each individually. •A user can generate their own segmented mappings to mask with unique values consistently, and define values which need to be preserved, such as dashes, or semantically rich values.	• Irreversible • Can be defined to maintain uniqueness	• SSN • Credit Card # • Primary/Foreign Keys • Account Numbers

ALGORITHMS

3. How to make friends with testing teams and show ROI at the same time

Challenge

- *“This is going to slow down and complicate my work.”*

Solution

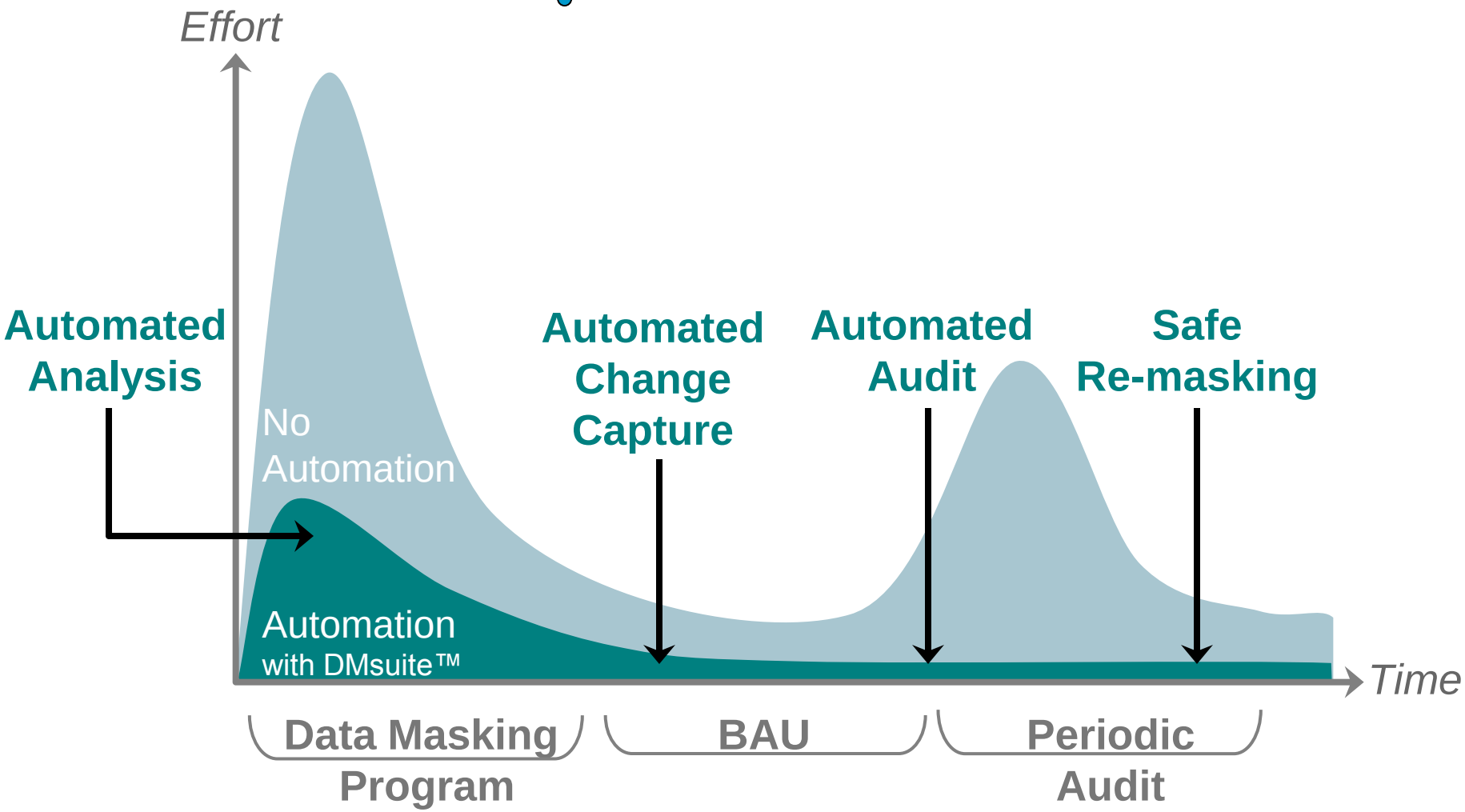
- Place more stringent controls on Production data.
- Employ automation wherever possible.

Benefit

- A side benefit of an automated approach is that it creates greater efficiency in maintaining data masking as well as in other areas.
- With a better understanding of how the application and its data, bug- and break-fixing becomes faster and easier.

A 'zero business value' project…

3. How to make friends with testing teams and show ROI at the same time



4. Showing results to ISOs, auditors, regulators, and sponsors

Challenge

- *"We did masking last year, so we're all set."*

Solution

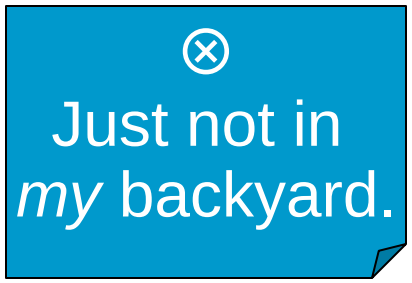
- Employ smart tools, process, automation and audit trails in the ongoing monitoring and periodic testing of masked environments.

Benefit

- Make it easy to keep and show that a masked environment is still 'clean.'

4. Showing results to ISOs, auditors, regulators, and sponsors

CERTIFY ENVIRONMENT IS STILL 'CLEAN'	START	END	OWNER
Participate in Typical Audit			
Identify applications and environment (calendar)			
Show Sensitive Data Inventory			
Show Data Context Diagram			
Show Data Provisioning Process Diagram			
Show Data Samples from environment			
Sign-off for Certification			
Run Profiler Test (if required)			
Run Profiler to detect presence of any sensitive data			
Perform typical activities (e.g. wait 1 work week)			
Run Profiler to detect presence of any sensitive data			
Show results before v.s after			
Research and fix any issues found and re-run test			
Sign-off for Certification			
Demonstrate Masking (if required)			
Create demo area			
Create mock Prod data (tip: use Auditor's name)			
Mask mock Prod data			
Show data before vs. after			
Document then delete demo			
Sign-off for Certification			



5. Where data masking fits into your overall information security framework of controls

Challenge

- *"We already don't just give 'anyone' access…"*

Solution

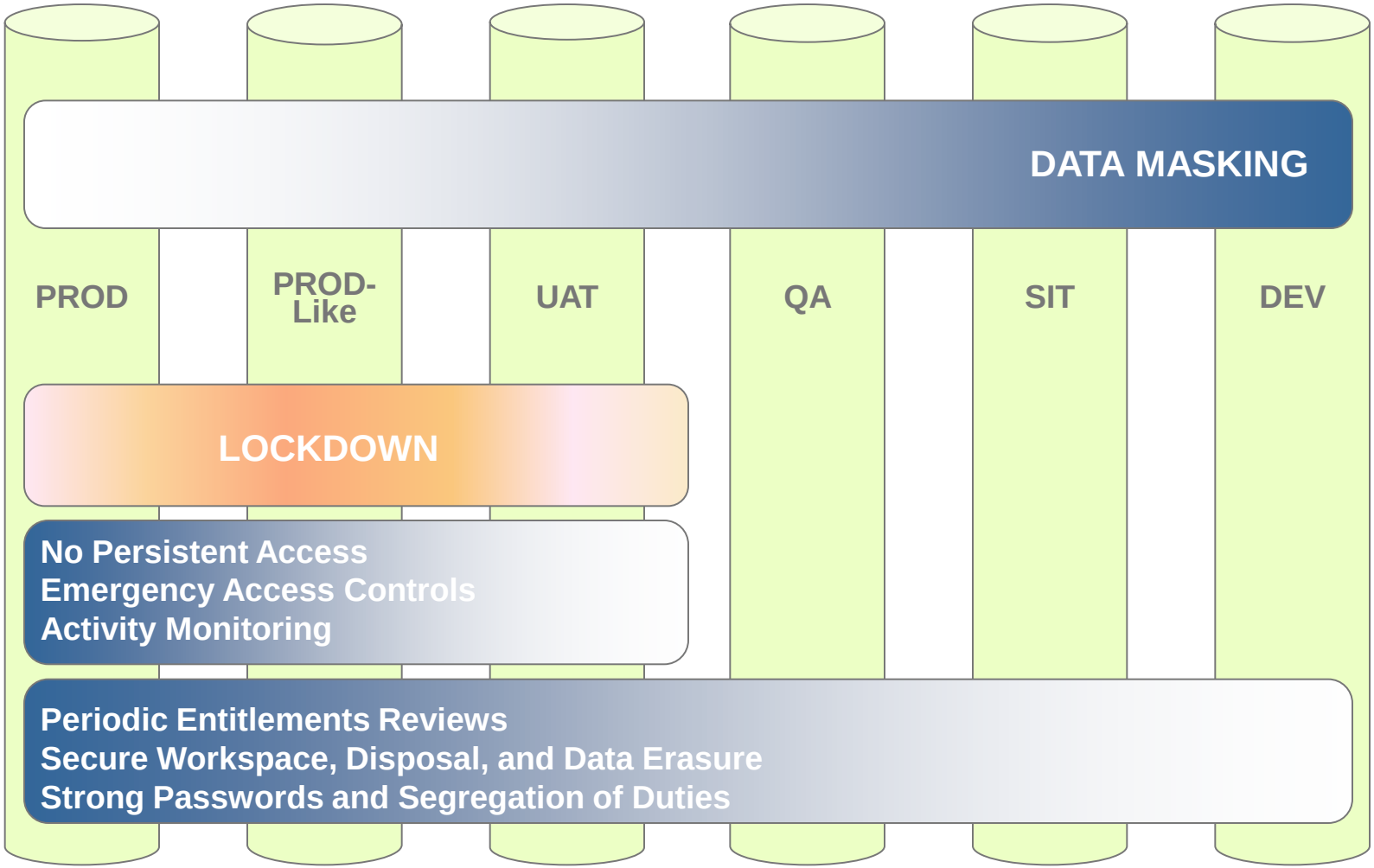
- Keep in mind that Data Masking is not a one-time event; it's one of several tools in your Data Security Toolkit.

Benefit

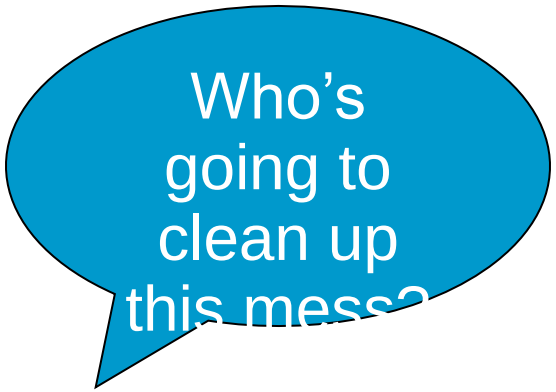
- Knowing when and where to employ data masking versus other data confidentiality controls (RBAC, lockdown, etc.) helps your organization avoid a slow-down in productivity.

⊗ Just not in *my* backyard.

5. Where data masking fits into your overall information security framework of controls



INFORMATION SECURITY CONTROLS



6. Integrating data masking into the application development lifecycle

Challenge

- *"How am I supposed to get my job done if I have to mask data at every step of the way?"*

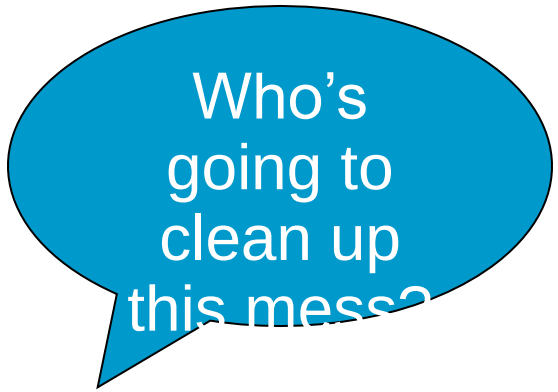
Solution

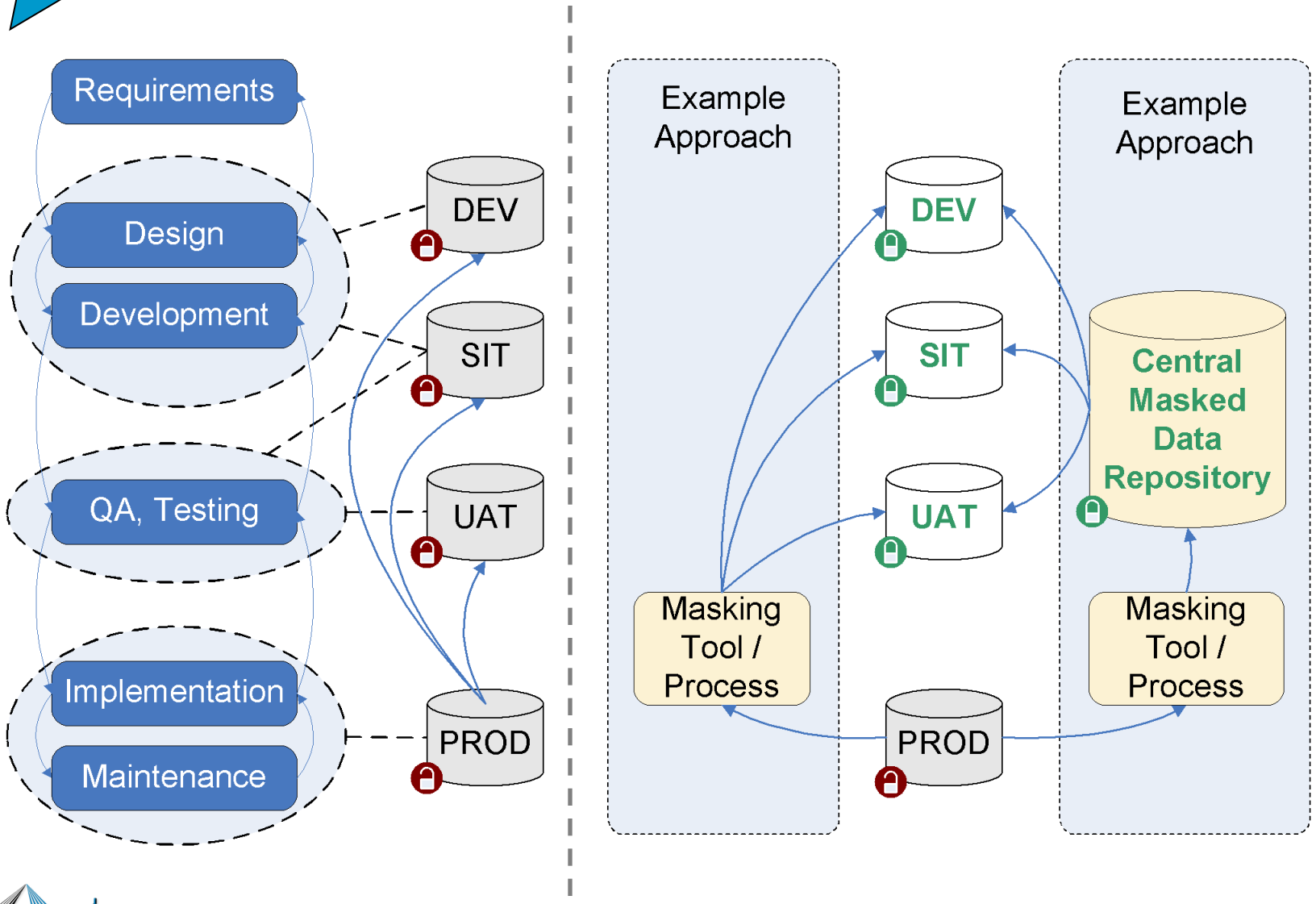
- Analyze each SDLC instance to determine the best potential 'in point' for data masking.

Benefit

- Data masking becomes part of the ongoing process, yet remains as behind-the-scenes as possible.

6. Integrating data masking into the application development lifecycle





SDLC INTEGRATION

7. Options for a shared center of excellence architecture for data masking

Challenge

- *"Sure this'll works for us, but our systems feed and receive data across businesses."*

Solution

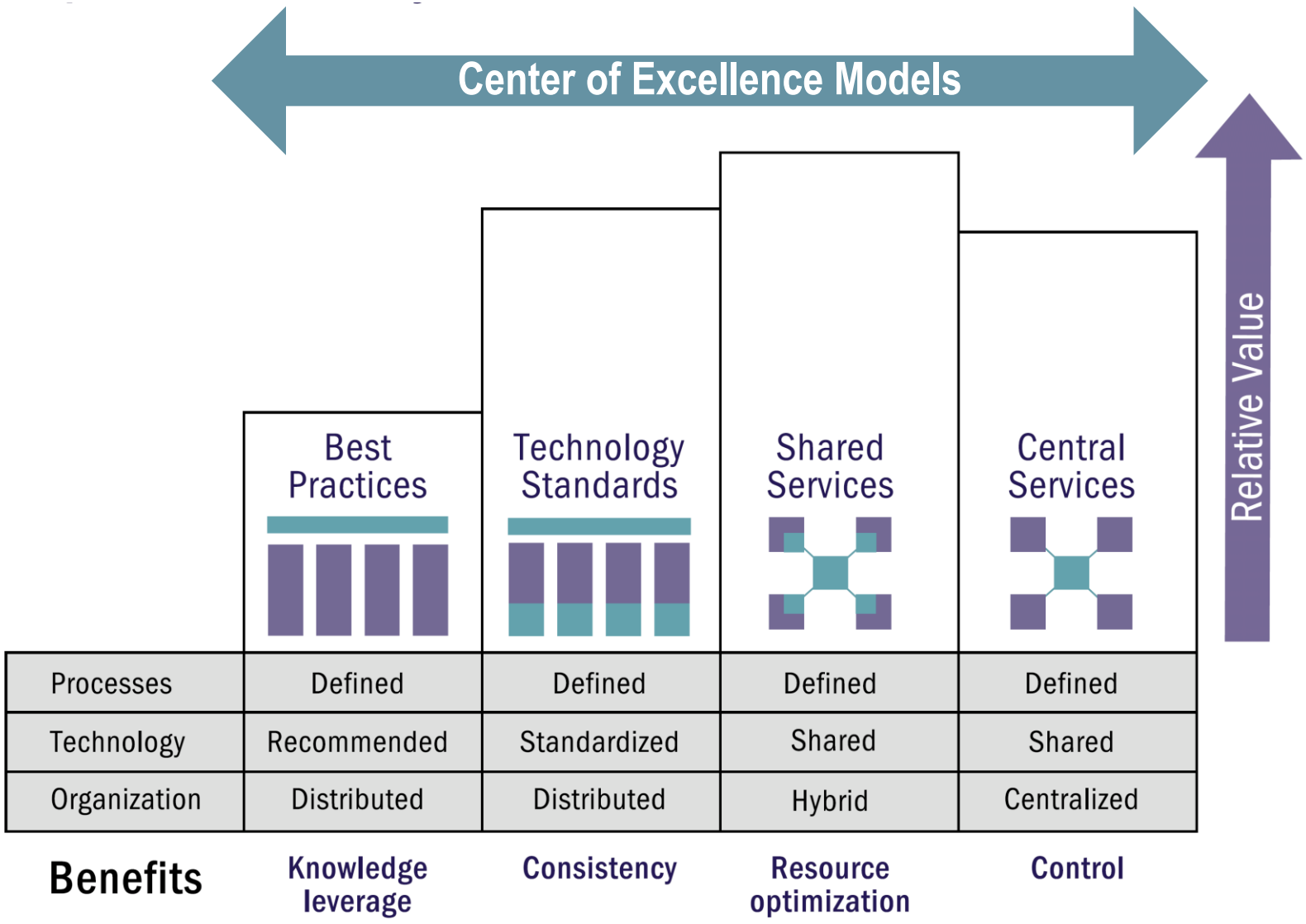
- Build for at least a basic-level Center of Excellence for Data Masking.

Benefit

- Achieve economies of scale by enabling your businesses to share knowledge about and resources for:
  - ✓ Process
  - ✓ Technology
  - ✓ Organization

Can't we all just get along?

7. Options for a shared center of excellence architecture for data masking



	Best Practices	Technology Standards	Shared Services	Central Services
Processes	Defined	Defined	Defined	Defined
Technology	Recommended	Standardized	Shared	Shared
Organization	Distributed	Distributed	Hybrid	Centralized
Benefits	Knowledge leverage	Consistency	Resource optimization	Control

Axis Technology, LLC & Data Masking

DMSuite™ - A robust, proprietary tool with data masking and self-service provisioning functionality. Straight out of the box, this software enables you to easily protect, provision and audit sensitive data from most known data sources (Oracle, DB2 (UDB & MF), Sybase, Informix, Teradata, MSSQL, MF Files (VSAM, QSAM), ADABAS, flat files (CVS, Excel, etc.)

Clients

A sampling of Axis Data Masking Projects

Morgan Stanley

- Data Masking with Informatica

Fidelity

- **Fidelity Brokerage(FBCT):** Mainframe and Unix-based applications
- **Fidelity CFIT**: Data Warehouse, Oracle Financials, and custom applications

New York Life Insurance

- Sensitive Data Assessment
- Data Masking Inventory management with DMsuite

State of MA

- Data Masking to comply with state regulations

State Street

- Data Masking to comply with state regulations

Bridging Strategy and Data

Axis Technology, LLC

Boston • New York • Dallas

185 Devonshire Street
Boston, MA 02110

www.AxisTechnologyLLC.com

THANK YOU.