Contents

Introduction xiii

Acknowledgments xv

About the Author xix

Chapter 1 Overview of Data Anonymization 1

Points to Ponder 1

PII 2

PHI 4

What Is Data Anonymization? 4

What Are the Drivers for Data Anonymization? 5

The Need to Protect Sensitive Data Handled as Part of Business 5

Increasing Instances of Insider Data Leakage, Misuse of Personal Data, and the Lure of Money for Mischievous Insiders 6

Astronomical Cost to the Business Due to Misuse of Personal Data 7

Risks Arising out of Operational Factors Such as Outsourcing and Partner Collaboration 8

Legal and Compliance Requirements 8

Will Procuring and Implementing a Data Anonymization Tool by Itself Ensure Protection of Privacy of Sensitive Data? 9

Ambiguity of Operational Aspects 10

Allowing the Same Users to Access Both Masked and Unmasked Environments 10

Lack of Buy-In from IT Application Developers, Testers, and End-Users 10

Compartmentalized Approach to Data Anonymization 11
Absence of Data Privacy Protection Policies or Weak Enforcement of Data Privacy Policies 11
Benefits of Data Anonymization Implementation 11
Conclusion 12
References 12

Part I Data Anonymization Program Sponsor's Guidebook

Chapter 2 Enterprise Data Privacy Governance Model 19
Points to Ponder 19
Chief Privacy Officer 20
Unit/Department Privacy Compliance Officers 22
The Steering Committee for Data Privacy Protection Initiatives 22
Management Representatives 23
Information Security and Risk Department Representatives 23
Representatives from the Departmental Security and Privacy Compliance Officers 24
Incident Response Team 24
The Role of the Employee in Privacy Protection 25
The Role of the CIO 26
Typical Ways Enterprises Enforce Privacy Policies 26
Conclusion 26

Chapter 3 Enterprise Data Classification Policy and Privacy Laws 29
Points to Ponder 29
Regulatory Compliance 30
Enterprise Data Classification 34
Points to Consider 36
Controls for Each Class of Enterprise Data 36
Conclusion 37

Chapter 4 Operational Processes, Guidelines, and Controls for Enterprise Data Privacy Protection 39
Points to Ponder 39
Privacy Incident Management 43
Planning for Incident Resolution 44
Preparation 45
Incident Capture 46
Incident Response 47
Post Incident Analysis 47
Guidelines and Best Practices 48
PII/PHI Collection Guidelines 48
Guidelines for Storage and Transmission of PII/PHI 49
PII/PHI Usage Guidelines 49

Guidelines for Storing PII/PHI on Portable Devices and Storage Devices 50
Guidelines for Staff 50
Conclusion 50
References 51

CHAPTER 5 THE DIFFERENT PHASES OF A DATA ANONYMIZATION PROGRAM 53
Points to Ponder 53
How Should I Go about the Enterprise Data Anonymization Program? 53
The Assessment Phase 54
Tool Evaluation and Solution Definition Phase 56
Data Anonymization Implementation Phase 56
Operations Phase or the Steady-State Phase 57
Food for Thought 58
When Should the Organization Invest in a Data Anonymization Exercise? 58
The Organization's Security Policies Mandate Authorization to Be Built into Every Application. Won't this Be Sufficient? Why is Data Anonymization Needed? 58
Is There a Business Case for a Data Anonymization Program in My Organization? 59
When Can a Data Anonymization Program Be Called a Successful One? 60
Why Should I Go for a Data Anonymization Tool When SQL Encryption Scripts Can Be Used to Anonymize Data? 61
Challenges with Using the SQL Encryption Scripts Approach for Data Anonymization 61
What Are the Benefits Provided by Data Masking Tools for Data Anonymization? 62
Why Is a Tool Evaluation Phase Needed? 62
Who Should Implement Data Anonymization? Should It Be the Tool Vendor, the IT Service Partner, External Consultants, or Internal Employees? 63
How Many Rounds of Testing Must Be Planned to Certify That Application Behavior Is Unchanged with Use of Anonymized Data? 64
Conclusion 64
Reference 65

CHAPTER 6 DEPARTMENTS INVOLVED IN ENTERPRISE DATA ANONYMIZATION PROGRAM 67
Points to Ponder 67
The Role of the Information Security and Risk Department 67
The Role of the Legal Department 68
The Role of Application Owners and Business Analysts 70

The Role of Administrators 70
The Role of the Project Management Office (PMO) 71
The Role of the Finance Department 71
Steering Committee 71
Conclusion 72

CHAPTER 7 PRIVACY METER—ASSESSING THE MATURITY OF DATA PRIVACY PROTECTION PRACTICES IN THE ORGANIZATION 75
Points to Ponder 75
Planning a Data Anonymization Implementation 78
Conclusion 79

CHAPTER 8 ENTERPRISE DATA ANONYMIZATION EXECUTION MODEL 83
Points to Ponder 83
Decentralized Model 84
Centralized Anonymization Setup 85
Shared Services Model 86
Conclusion 87

CHAPTER 9 TOOLS AND TECHNOLOGY 89
Points to Ponder 89
Shortlisting Tools for Evaluation 91
Tool Evaluation and Selection 92
Functional Capabilities 92
Technical Capabilities 96
Operational Capabilities 99
Financial Parameters 99
Scoring Criteria for Evaluation 101
Conclusion 101

CHAPTER 10 ANONYMIZATION IMPLEMENTATION—ACTIVITIES AND EFFORT 103
Points to Ponder 103
Anonymization Implementation Activities for an Application 104
Application Anonymization Analysis and Design 104
Anonymization Environment Setup 105
Application Anonymization Configuration and Build 105
Anonymized Application Testing 105
Complexity Criteria 105
Application Characteristics 106
Environment Dependencies 106
Arriving at an Effort Estimation Model 107
Case Study 108
Context 108
Estimation Approach 109
Application Characteristics for LOANADM 110

Arriving at a Ball Park Estimate 110
Conclusion 111

CHAPTER 11 THE NEXT WAVE OF DATA PRIVACY CHALLENGES 113

PART II DATA ANONYMIZATION PRACTITIONER'S GUIDE

CHAPTER 12 DATA ANONYMIZATION PATTERNS 119
Points to Ponder 119
Pattern Overview 119
Conclusion 121

CHAPTER 13 DATA STATE ANONYMIZATION PATTERNS 123
Points to Ponder 123
Principles of Anonymization 123
Static Masking Patterns 124
EAL Pattern (Extract-Anonymize-Load Pattern) 125
ELA Pattern (Extract-Load-Anonymize Pattern) 125
Data Subsetting 126
Dynamic Masking 128
Dynamic Masking Patterns 128
Interception Pattern 129
When Should Interception Patterns be Selected and on What Basis? 130
Challenges Faced When Implementing Dynamic Masking Leveraging Interception Patterns 132
Invocation Pattern 132
Application of Dynamic Masking Patterns 133
Dynamic Masking versus Static Masking 133
Conclusion 134

CHAPTER 14 ANONYMIZATION ENVIRONMENT PATTERNS 137
Points to Ponder 137
Application Environments in an Enterprise 137
Testing Environments 139
Standalone Environment 140
Integration Environment 141
Automated Integration Test Environment 144
Scaled-Down Integration Test Environment 148
Conclusion 150

CHAPTER 15 DATA FLOW PATTERNS ACROSS ENVIRONMENTS 153
Points to Ponder 153
Flow of Data from Production Environment Databases to Nonproduction Environment Databases 153
Controls Followed 155

Movement of Anonymized Files from Production Environment to Nonproduction Environments 155
Controls 157
Masked Environment for Integration Testing—Case Study 157
Objectives of the Anonymization Solution 158
Key Anonymization Solution Principles 158
Solution Implementation 159
Anonymization Environment Design 160
Anonymization Solution 161
Anonymization Solution for the Regression Test/ Functional Testing Environment 163
Anonymization Solution for an Integration Testing Environment 163
Anonymization Solution for UAT Environment 164
Anonymization Solution for Preproduction Environment 164
Anonymization Solution for Performance Test Environment 165
Anonymization Solution for Training Environment 166
Reusing the Anonymization Infrastructure across the Various Environments 166
Conclusion 169
Anonymization Environment Design 169

CHAPTER 16 DATA ANONYMIZATION TECHNIQUES 171
Points to Ponder 171
Basic Anonymization Techniques 172
Substitution 172
Shuffling 174
Number Variance 176
Date Variance 177
Character Masking 181
Cryptographic Techniques 182
Partial Sensitivity and Partial Masking 185
Masking Based on External Dependancy 185
Auxiliary Anonymization Techniques 186
Alternate Classification of Data Anonymization Techniques 189
Leveraging Data Anonymization Techniques 190
Case Study 191
Input File Structure 191
AppTable Structure 191
Output File Structure 194
Solution 194
Conclusion 195
Data Anonymization Mandatory and Optional Principles 196
Reference 196

CHAPTER 17 DATA ANONYMIZATION IMPLEMENTATION 197
Points to Ponder 197

Prerequisites before Starting Anonymization Implementation Activities 199
Sensitivity Definition Readiness—What Is Considered Sensitive Data by the Organization? 199
Sensitive Data Discovery—Where Do Sensitive Data Exist? 200
Application Architecture Analysis 200
Application Sensitivity Analysis 202
What Is the Sensitivity Level and How Do We Prioritize Sensitive Fields for Treatment? 203
Case Study 204
Anonymization Design Phase 208
Choosing an Anomymization Technique for Anonymization of Each Sensitive Field 208
Choosing a Pattern for Anonymization 209
Anonymization Implementation, Testing, and Rollout Phase 211
Anonymization Controls 212
Anonymization Operations 213
Incorporation of Privacy Protection Procedures as Part of Software Development Life Cycle and Application Life Cycle for New Applications 214
Impact on SDLC Team 216
Challenges Faced as Part of Any Data Anonymization Implementation 216
General Challenges 216
Functional, Technical, and Process Challenges 217
People Challenges 219
Best Practices to Ensure Success of Anonymization Projects 220
Creation of an Enterprise-Sensitive Data Repository 220
Engaging Multiple Stakeholders Early 220
Incorporating Privacy Protection Practices into SDLC and Application Life Cycle 220
Conclusion 221
References 221
APPENDIX A: GLOSSARY 223
INDEX 229