



# DOCUMENT ENGINEERING

**ANALYZING AND DESIGNING  
DOCUMENTS FOR BUSINESS  
INFORMATICS AND WEB SERVICES**

**ROBERT J. GLUSHKO  
AND TIM McGRATH**



# CONTENTS

---

PREFACE	xv
FOREWORD	xxi

## **PART I** **INTRODUCTION**

### **CHAPTER ONE** **INTRODUCTION TO DOCUMENT ENGINEERING**

1.0 INTRODUCTION	4
1.1 A SIMPLE BUSINESS TRANSACTION?	6
1.2 THE EXTENDED OR VIRTUAL ENTERPRISE	8
1.3 IT'S ALL ABOUT EXCHANGING DOCUMENTS	9
1.3.1 THE DOCUMENT TYPE SPECTRUM	10
1.3.2 DOCUMENT EXCHANGE AS A BUILDING BLOCK IN BUSINESS MODEL PATTERNS	11
1.3.3 DOCUMENT EXCHANGE AS LOOSE COUPLING	13
1.4 UNDERSTANDING THE MEANING OF DOCUMENTS	14
1.4.1 INCOMPATIBLE INFORMATION MODELS	16
1.4.2 STANDARD INFORMATION MODELS	17
1.5 XML AS AN ENABLING TECHNOLOGY	17
1.6 USING XML-ENCODED MODELS TO DESIGN AND DRIVE APPLICATIONS	18
1.6.1 DOCUMENT MODELS AS INTERFACES	20
1.6.2 MODELS OF BUSINESS PROCESSES	22
1.6.3 WEB SERVICES AND SERVICE ORIENTED ARCHITECTURES	23
1.7 DOCUMENT SPECIFICATIONS AND DOCUMENT ENGINEERING	26
1.8 DOCUMENT ENGINEERING—A NEW AND SYNTHETIC DISCIPLINE	27
1.8.1 BUSINESS PROCESS ANALYSIS	28
1.8.2 TASK ANALYSIS	29
1.8.3 DOCUMENT ANALYSIS	29
1.8.4 DATA ANALYSIS	30
1.8.5 UNIFICATION IN DOCUMENT ENGINEERING	31
1.9 THE DOCUMENT ENGINEERING APPROACH	32
1.10 KEY POINTS IN CHAPTER ONE	37

## **PART II**

### **CHAPTER TWO**

2.0 INTRODUCTION	42
2.1 FROM HTML TO XML	43
2.1.1 THE BROWSER WAR	46
2.1.2 FROM THE WEB FOR EYES TO THE WEB FOR COMPUTERS	47
2.2 XML'S BIG IDEAS	48
2.3 CREATION OF NEW SETS OF TAGS FOR DOMAIN-SPECIFIC CONTENT	49
2.4 SEPARATION OF CONTENT AND PRESENTATION	51
2.5 DEFINITION OF DOCUMENT TYPES	53
2.5.1 DOCUMENTS AS IMPLEMENTATION MODELS	54
2.5.2 XML SCHEMAS	55
2.5.3. SCHEMA LANGUAGES	56
2.5.4 RULES THAT SCHEMA LANGUAGES CAN'T REPRESENT	61
2.6 VALIDATION	62
2.7 CONVERSION AND TRANSFORMATION	64
2.7.1 CONVERSION TO XML	64
2.7.2 TRANSFORMATION FROM XML	67
2.7.3 WHERE TO TRANSFORM	69
2.8 KEY POINTS IN CHAPTER TWO	72

### **CHAPTER THREE**

### **MODELS, PATTERNS, AND REUSE**

3.0 INTRODUCTION	74
3.1 MODELS	75
3.2 ADAPTING THE CLASSIC MODELING APPROACH TO DOCUMENTS	76
3.2.1 EXTERNAL VIEWS: INSTANCES OF DOCUMENT IMPLEMENTATIONS	78
3.2.2 PHYSICAL VIEWS: DOCUMENT IMPLEMENTATION MODELS (OR SCHEMAS)	79
3.2.3 CONCEPTUAL VIEWS: DOCUMENT COMPONENT AND ASSEMBLY MODELS	81
3.3 THE MODEL MATRIX	86
3.3.1 METADATA AND METAMODELS	88
3.3.2 METAMODELS FOR PROCESSES	89

## **FOUNDATIONS**

### **XML FOUNDATIONS**



5.7 FROM PROPRIETARY TO STANDARD MODELS	163
5.8 TOWARD MODELS WITH REUSABLE COMPONENTS	164
5.9 KEY POINTS IN CHAPTER FIVE	168

## **CHAPTER SIX**

## **WHEN MODELS DON'T MATCH: THE INTEROPERABILITY CHALLENGE**

6.0 INTRODUCTION	172
6.1 THE INTEROPERABILITY CHALLENGE	175
6.1.1 THE INTEROPERABILITY TARGET	176
6.1.2 RECOGNIZING EQUIVALENCE	180
6.2 CONTENT CONFLICTS	181
6.3 ENCODING CONFLICTS	184
6.3.1 LANGUAGE CONFLICTS	184
6.3.2 GRAMMATICAL CONFLICTS	186
6.4 STRUCTURAL CONFLICTS	187
6.4.1 DOCUMENT ASSEMBLY CONFLICTS	188
6.4.2 COMPONENT ASSEMBLY CONFLICTS	189
6.4.3 COMPONENT GRANULARITY CONFLICTS	191
6.5 SEMANTIC CONFLICTS	193
6.5.1 FUNCTIONAL DEPENDENCY CONFLICTS	194
6.5.2 VOCABULARY CONFLICTS	195
6.5.3 CONTEXTUAL CONFLICTS	198
6.6 MOTIVATING THE DOCUMENT ENGINEERING APPROACH	200
6.7 KEY POINTS IN CHAPTER SIX	203

## **PART III THE DOCUMENT ENGINEERING APPROACH**

### **CHAPTER SEVEN THE DOCUMENT ENGINEERING APPROACH**

7.0 INTRODUCTION	208
7.1 AN APPROACH, NOT A METHODOLOGY	208
7.1.1 MODELING METHODOLOGIES	210
7.2 THE DOCUMENT ENGINEERING APPROACH	211
7.2.1 UNIFIED APPROACH TO MODELING	212
7.2.2 ARTIFACT-FOCUSED VIEW OF MODELING	214
7.2.3 THE MODELING PHASES, TASKS, AND ARTIFACTS	215
7.2.4 IMPLEMENTING MODELS IN APPLICATIONS	219

7.3 ANALYZING THE CONTEXT OF USE	221
7.3.1 REQUIREMENTS	221
7.3.2 PATTERNS OF CONTEXT	223
7.3.3 SCOPE OF CONTEXT	224
7.4 ANALYZING BUSINESS PROCESSES	224
7.5 DESIGNING BUSINESS PROCESSES	226
7.6 ANALYZING DOCUMENTS	226
7.6.1 <i>CREATE THE DOCUMENT INVENTORY</i>	227
7.6.2 <i>SAMPLE THE DOCUMENT INVENTORY</i>	228
7.7 ANALYZING DOCUMENT COMPONENTS	228
7.7.1 <i>HARVEST THE COMPONENTS</i>	229
7.7.2 <i>NAME THE CONTENT COMPONENTS</i>	231
7.7.3 <i>CONSOLIDATE THE CANDIDATE COMPONENTS</i>	232
7.8 ASSEMBLING DOCUMENT COMPONENTS	232
7.8.1 <i>FORMALIZE THE COMPONENT MODEL</i>	233
7.8.2 <i>ASSOCIATIONS BETWEEN STRUCTURES</i>	234
7.8.3 <i>REFINE THE COMPONENT NAMES</i>	234
7.9 ASSEMBLING DOCUMENT MODELS	235
7.10 IMPLEMENTING MODELS	236
7.10.1 <i>ENCODING DOCUMENT IMPLEMENTATION MODELS</i>	237
7.10.2 <i>ENCODING BUSINESS PROCESS IMPLEMENTATION MODELS</i>	238
7.11 SUMMARY OF MODELING PHASES AND ARTIFACTS	239
7.12 KEY POINTS IN CHAPTER SEVEN	241

## **CHAPTER EIGHT**

## **ANALYZING THE CONTEXT OF USE**

8.0 INTRODUCTION	244
8.1 UNDERSTANDING DOCUMENT AND PROCESS REQUIREMENTS	247
8.1.1 <i>STRATEGIC AND TACTICAL REQUIREMENTS</i>	248
8.1.2 <i>SOURCES OF REQUIREMENTS</i>	250
8.1.3 <i>GENERIC REQUIREMENTS</i>	252
8.2 CONTEXT AND REQUIREMENTS	254
8.3 EXPRESSING REQUIREMENTS AS RULES	259
8.3.1 <i>USAGE REQUIREMENTS</i>	261
8.3.2 <i>STRUCTURAL REQUIREMENTS</i>	262
8.3.3 <i>SEMANTIC REQUIREMENTS</i>	263
8.3.4 <i>PRESENTATION REQUIREMENTS</i>	264



## **CHAPTER TEN**

## **DESIGNING BUSINESS PROCESSES WITH PATTERNS**

10.0 INTRODUCTION	318
10.1 WHY WE USE PATTERNS IN PROCESS MODELS	318
10.2 HOW WE USE PATTERNS IN PROCESS MODELS	320
10.3 PATTERNS AND THE MODEL MATRIX	321
10.4 IDENTIFYING CANDIDATE DESIGN PATTERNS	324
10.4.1 GENERALIZING PATTERNS	327
10.4.2 VARYING THE GRANULARITY OF PATTERNS	329
10.4.3 COMBINING PATTERNS	332
10.4.4 USING IMPLEMENTATIONS AS PATTERNS	334
10.5 CHOOSING APPROPRIATE PATTERNS	337
10.5.1 VALIDATING REQUIREMENTS VS. DISCOVERING THEM	337
10.5.2 REINFORCING CONTEXTS WITH PATTERNS	338
10.5.3 APPLYING PATTERNS TO ACHIEVE INSIGHT	340
10.6 ADAPTING PATTERNS	341
10.7 INSTANTIATING PATTERNS TO CREATE NEW MODELS	343
10.7.1 INSTANTIATING ROLES	343
10.7.2 CONFIGURING COLLABORATION AND TRANSACTION PROPERTIES	344
10.8 USING PATTERNS TO SUGGEST INFORMATION COMPONENTS AND DOCUMENTS	347
10.8.1 KEY INFORMATION COMPONENTS	347
10.8.2 THE DOCUMENT CHECKLIST	349
10.9 KEY POINTS IN CHAPTER TEN	351

## **CHAPTER ELEVEN**

## **ANALYZING DOCUMENTS**

11.0 INTRODUCTION	354
11.1 WHAT ARE DOCUMENTS?	355
11.2 CREATING THE INVENTORY	356
11.2.1 DOCUMENT ARCHAEOLOGY AND ANTHROPOLOGY	357
11.2.2 UNDERSTANDING THE ORGANIZATION	359
11.2.3 GENERIC INVENTORY PROCEDURES AND QUESTIONS	360
11.3 SAMPLING THE INVENTORY	364
11.3.1 SAMPLING BASED ON DOCUMENT CHARACTERISTICS	364
11.3.2 SAMPLING BASED ON OTHER CONSIDERATIONS	369
11.4 KEY POINTS IN CHAPTER ELEVEN	371

## **CHAPTER TWELVE                    ANALYZING DOCUMENT COMPONENTS**

12.0 INTRODUCTION	374
12.1 HARVESTING COMPONENTS	374
12.1.1 FINDING COMPONENTS IN TRANSACTIONAL DOCUMENTS	376
12.1.2 FINDING COMPONENTS IN NARRATIVE DOCUMENTS	378
12.1.3 FINDING STRUCTURES FOR COMPONENTS	380
12.1.4 FINDING SEMANTIC CONTENT IN PRESENTATION COMPONENTS	380
12.1.5 FINDING MEANINGFUL CONTENT IN STRUCTURAL COMPONENTS	386
12.1.6 THE TROUBLE WITH TABLES	389
12.1.7 PRESERVING PRESENTATION COMPONENTS	393
12.1.8 ANALYZING SETS OF POSSIBLE VALUES	393
12.1.9 CODE SETS	399
12.1.10 IDENTIFIERS	402
12.1.11 NAMING COMPONENTS	402
12.2 CONSOLIDATING COMPONENTS	406
12.2.1 CREATING A CONSOLIDATED TABLE OF CONTENT COMPONENTS	406
12.2.2 ENSURING SEMANTIC UNIQUENESS	412
12.3 KEY POINTS IN CHAPTER TWELVE	416

## **CHAPTER THIRTEEN                ASSEMBLING DOCUMENT COMPONENTS**

13.0 INTRODUCTION	420
13.1 ASSOCIATIONS BETWEEN COMPONENTS	421
13.2 DOCUMENT COMPONENT MODELS	422
13.2.1 RULES IN NARRATIVE DOCUMENT CONTEXTS	423
13.2.2 RULES IN TRANSACTIONAL CONTEXTS	426
13.3 METHODS FOR AGGREGATING COMPONENTS	427
13.3.1 THE CLASSICAL DOCUMENT ANALYSIS APPROACH	427
13.3.2 THE CLASSICAL DATA ANALYSIS APPROACH	429
13.4 APPLYING NORMALIZATION TO DOCUMENT ENGINEERING	429
13.4.1 FUNCTIONAL DEPENDENCY	430
13.4.2 ESSENTIALITY	431
13.4.3 THE NORMAL FORMS	434
13.4.4 IDENTIFYING PRIMARY KEYS	434
13.4.5 RECURSIVE ASSOCIATIONS	441
13.4.6 MULTIVALUE DEPENDENCIES	443



15.3 MODEL BASED APPLICATIONS IN DOCUMENT ENGINEERING	510
15.3.1 E-FORM APPLICATIONS	511
15.3.2 E-BOOK APPLICATIONS	521
15.3.3 SINGLE SOURCE PUBLISHING AND PORTAL APPLICATIONS	522
15.3.4 BUSINESS-TO-BUSINESS DOCUMENT EXCHANGES	525
15.3.5 APPLICATIONS WITH INTERMEDIARY PLATFORMS	528
15.3.6 COMPOSITE SERVICES	531
15.3.7 SEMANTIC WEB AND SEMANTIC WEB SERVICE APPLICATIONS	532
15.4 IMPLEMENTING MODELS IN APPLICATIONS: THE FUTURE	535
15.5 KEY POINTS IN CHAPTER FIFTEEN	536

## **CHAPTER SIXTEEN**

## **MANAGEMENT AND STRATEGY**

16.0 INTRODUCTION	540
16.1 ORGANIZATIONAL MATURITY	542
16.1.1 MOTIVATING CAPABILITY ASSESSMENT	543
16.1.2 THE CAPABILITY MATURITY MODEL	543
16.1.3 THE DOCUMENT ENGINEERING CAPABILITY MATURITY MODEL	545
16.1.4 CONDUCTING A CAPABILITY ASSESSMENT	550
16.2 BUSINESS OBJECTIVES	554
16.2.1 MAKING A BUSINESS CASE	555
16.2.2 A SAMPLE OF PROJECT JUSTIFICATIONS	558
16.2.3 A SAMPLE OF PROJECT RISKS	566
16.3 KEY POINTS IN CHAPTER SIXTEEN	571

## **PART IV**

## **THE END OF THE BEGINNING**

### **CHAPTER SEVENTEEN**

### **EPILOGUE**

17.0 INTRODUCTION	576
17.1 WHEN DISCIPLINES COLLIDE	576
17.2 THE BUSINESS OF DOCUMENT ENGINEERING	577
17.3 THE SUCCESSFUL DOCUMENT ENGINEER	578

NOTES	579
GLOSSARY	621
INDEX	647