CMIS and Apache Chemistry in Action

FLORIAN MÜLLER
JAY BROWN
JEFF POTTS

MANNING
SHELTER ISLAND

Licensed to Arti Pandey <arti.pandey@dreamtechpress.com>

For online information and ordering of this and other Manning books, please visit www.manning.com. The publisher offers discounts on this book when ordered in quantity. For more information, please contact

Special Sales Department
Manning Publications Co.
20 Baldwin Road
PO Box 261
Shelter Island, NY 11964
Email: orders@manning.com

©2013 by Manning Publications Co. All rights reserved.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by means electronic, mechanical, photocopying, or otherwise, without prior written permission of the publisher.

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in the book, and Manning Publications was aware of a trademark claim, the designations have been printed in initial caps or all caps.

♾ Recognizing the importance of preserving what has been written, it is Manning's policy to have the books we publish printed on acid-free paper, and we exert our best efforts to that end. Recognizing also our responsibility to conserve the resources of our planet, Manning books are printed on paper that is at least 15 percent recycled and processed without the use of elemental chlorine.

Manning Publications Co.
20 Baldwin Road
PO Box 261
Shelter Island, NY 11964

Development editor: Karen G. Miller
Technical proofreader: David Caruana
Copyeditors: Benjamin Berg, Andy Carroll
Proofreader: Katie Tennant
Typesetter: Dottie Marsico
Cover designer: Marija Tudor

ISBN 9781617291159
Printed in the United States of America
1 2 3 4 5 6 7 8 9 10 – MAL – 18 17 16 15 14 13

Licensed to Arti Pandey <arti.pandey@dreamtechpress.com>

brief contents

PART 1 UNDERSTANDING CMIS 1

1 ■ Introducing CMIS 3
2 ■ Exploring the CMIS domain model 19
3 ■ Creating, updating, and deleting objects with CMIS 39
4 ■ CMIS metadata: types and properties 58
5 ■ Query 83

PART 2 HANDS-ON CMIS CLIENT DEVELOPMENT..... 115

6 ■ Meet your new project: The Blend 117
7 ■ The Blend: read and query functionality 150
8 ■ The Blend: create, update, and delete functionality 193
9 ■ Using other client libraries 235
10 ■ Building mobile apps with CMIS 277

PART 3 ADVANCED TOPICS 313

11 ■ CMIS bindings 315
12 ■ Security and control 339
13 ■ Performance 354
14 ■ Building a CMIS server 368

Licensed to Arti Pandey <arti.pandey@dreamtechpress.com>

Licensed to Arti Pandey <arti.pandey@dreamtechpress.com>

contents

foreword by R.J. Howarth xv
foreword by J. Newton xviii
preface xxi
acknowledgments xxiii
about this book xxv
about the authors xxvii
about the cover illustration xxix

PART 1 UNDERSTANDING CMIS..... 1

1 ***Introducing CMIS* 3**

1.1 What is CMIS? 3

About the specification 6 ▪ *What does CMIS do?* 6
Where is CMIS being adopted? 8

1.2 Setting up a CMIS test environment 9

Requirements 10 ▪ *Installing the OpenCMIS InMemory Repository web application* 10 ▪ *Installing the CMIS Workbench* 11

1.3 Writing your first CMIS code using Groovy 12

Connecting to the repository 12 ▪ *Try it—browse the repository using the CMIS Workbench* 13 ▪ *Try it—run CMIS code in the CMIS Workbench Groovy console* 14

Licensed to Arti Pandey <arti.pandey@dreamtechpress.com>

1.4 CMIS considerations 16

Understanding the limitations of CMIS 16 ▪ Comparing CMIS to the Java Content Repository (JCR) API 17

1.5 Summary 18

2 *Exploring the CMIS domain model 19*

2.1 The CMIS service 20

The role of the CMIS service 21 ▪ Bindings: what does a CMIS service look like? 21

2.2 Repository—the CMIS database 22

Repository info and capabilities 23 ▪ Capabilities across different repository vendors 25 ▪ Try it—retrieve the repository info 25

2.3 Folders 26

The role of folders 27 ▪ Try it—folder navigation 28

2.4 Documents 29

The role of documents 30 ▪ Properties 31 ▪ Try it—list a document's properties 34 ▪ Content streams 35 ▪ Try it—retrieve a document's content stream 36

2.5 The item object type (version 1.1) 37

2.6 Summary 38

3 *Creating, updating, and deleting objects with CMIS 39*

3.1 Creating objects 40

Requirements for creating an object 40 ▪ Try it—create a folder 40 ▪ Things to think about when creating folders 41 Try it—create a document 42 ▪ Things to think about when creating documents 45

3.2 Updating objects 46

Try it—rename a document or a folder 47 ▪ Try it—update the content stream 47 ▪ Understanding versioning 50 ▪ Try it—upload a new version of a document 52

3.3 Deleting objects 55

Requirements for deleting objects 55 ▪ Try it—delete an object 55 Things to think about when deleting objects 57

3.4 Summary 57

Licensed to Arti Pandey <arti.pandey@dreamtechpress.com>

4 *CMIS metadata: types and properties 58*

4.1 What is metadata and why do we need it? 59

4.2 Metadata in CMIS 60

Type definitions are hierarchical and attributes are inherited 61
Try it—view the types and property definitions using Workbench 61

4.3 Type collections and hierarchies 63

Try it—traversing the type hierarchy 64 ▪ Try it—examining property definitions on types 66 ▪ Constraints on property definitions 69 ▪ Try it—examining constraints on property definitions 70 ▪ Attribute and attribute value inheritance 72

4.4 CMIS 1.1 metadata features 72

Type mutability 72 ▪ Secondary types 80

4.5 Summary 82

5 *Query 83*

5.1 Query: a familiar face on search 84

Prerequisite for this chapter: SQL basics 84 ▪ Exercises in this chapter and the InMemory server 84

5.2 Introduction to the CMIS Query language 85

Reviewing clauses of the SELECT statement 85 ▪ Checking Query capabilities on a service 86 ▪ Try it—checking the Query capabilities of a CMIS service 87 ▪ Try it—your first CMIS Query 87 ▪ Try it—running a query from code 89 ▪ Checking query-related attributes for properties 91 ▪ Search scope 92

5.3 Components of a query 93

The SELECT clause 94 ▪ WHERE clause 95 ▪ Ordering and limiting query results 102 ▪ Joins and determining repository support 103

5.4 CMIS SQL extension functions 105

CONTAINS(): full-text search 105 ▪ Score() 110
Navigational functions 111

5.5 Summary 113

Licensed to Arti Pandey <arti.pandey@dreamtechpress.com>

PART 2 HANDS-ON CMIS CLIENT DEVELOPMENT 115

6 *Meet your new project: The Blend 117*

6.1 Understanding the business requirements and technical approach 118

Business requirements 118 ▪ Establishing the technical design 119

6.2 Walking through the finished product 125

6.3 Setting up the development environment 131

6.4 Configuring the InMemory server 135

6.5 Taking first steps with The Blend 137

Setting up the Eclipse project 137 ▪ Creating a session factory 139 ▪ Creating the servlets 140 ▪ Creating the JSPs 145 ▪ Try it—testing The Blend 147

6.6 Summary 148

7 *The Blend: read and query functionality 150*

7.1 Building a browse page 151

Preparing the HTML part of the browse page 152 ▪ Getting the folder object 153 ▪ Taking advantage of the OperationContext 155 ▪ Getting the folder children 156 Paging 158 ▪ Getting the folder parent 160 ▪ Assembling the browse page 160

7.2 Building a document page 165

Preparing the HTML part of the document page 166 ▪ Retrieving documents 167 ▪ Assembling the document page 174 The download servlet 177 ▪ Adding the version series to the document page 180

7.3 Building a query page 183

Ways to query: there be three 184 ▪ Assembling the search page 189 ▪ Accessing and traversing relationships 191

7.4 Summary 192

8 *The Blend: create, update, and delete functionality 193*

8.1 Creating folders 193

Two ways to create folders 194 ▪ Create folder: doPost() 196 Enumerating the creatable folder types 198

Licensed to Arti Pandey <arti.pandey@dreamtechpress.com>

8.2 Creating documents 200

Creating doGet() and doPost() for document creation 201
Performing file uploads 203

8.3 Updating properties 208

Concurrent access and locking 210 ▪ Properties from CMIS 1.1 secondary types 212

8.4 Updating and deleting content 214

Deleting content 214 ▪ Replacing content 215 ▪ Appending content 216

8.5 Versioning 217

Creating a new version 218 ▪ The checkIn() method 220

8.6 Copying documents 223

8.7 Moving objects 226

8.8 Deleting objects 230

Deleting documents 231 ▪ Deleting folders 232

8.9 Summary 234

9 *Using other client libraries 235*

9.1 Working with other client libraries 236

Common client libraries 236

9.2 Coding in .NET with DotCMIS 237

Comparing DotCMIS and OpenCMIS 237 ▪ Getting started with DotCMIS 238 ▪ Try it—building a web part with .NET and CMIS to browse The Blend 243 ▪ Using SharePoint as a CMIS repository 248 ▪ Connecting to SharePoint 250

9.3 Coding in Python with cmislib 251

Comparing cmislib and OpenCMIS 253 ▪ Installing cmislib 254 ▪ Connecting to a CMIS repository using the interactive shell 254 ▪ Using cmislib to synchronize objects between two CMIS repositories 260

9.4 Apache Chemistry PHP API 267

Installing the PHP Client 268 ▪ About the PHP Client library 268 ▪ PHP Client architecture 268 ▪ Differences between OpenCMIS and the PHP Client 270 ▪ Using PHP to browse The Blend 272

9.5 Summary 276

Licensed to Arti Pandey <arti.pandey@dreamtechpress.com>

10 *Building mobile apps with CMIS 277*

10.1 Writing mobile apps with OpenCMIS for Android 278

Android and CMIS 278 ▪ Setting up an Android environment 279 ▪ Writing your first Android CMIS application 282 ▪ Try it—writing an Android application for The Blend 286

10.2 Writing iOS apps with ObjectiveCMIS 292

What is ObjectiveCMIS? 292 ▪ Comparing ObjectiveCMIS with OpenCMIS 293 ▪ Getting started with ObjectiveCMIS 294 Using ObjectiveCMIS 302 ▪ Try it—writing an iOS application to capture new tracks for The Blend 305

10.3 Summary 310

PART 3 ADVANCED TOPICS 313

11 *CMIS bindings 315*

11.1 CMIS binding overview 316

The RESTful trend 316 ▪ The need for JavaScript support 316 Capturing CMIS traffic for inspection 317 ▪ Try it—tracing requests from part 1 317

11.2 A close look at the three bindings 318

The Web Services binding 318 ▪ The AtomPub binding 323 The Browser binding 329

11.3 CMIS schemas and schema extensions 334

XML schema 335

11.4 The OpenCMIS low-level API 336

Reasons to use the low-level API 337

11.5 Summary 337

12 *Security and control 339*

12.1 General security considerations 339

Cross-site scripting (XSS) attacks 340 ▪ Cross-site request forgery (CSRF) attacks 341

12.2 Authentication 341

Cookies 342 ▪ AuthenticationProvider interface 342 Example of an authentication provider 343

Licensed to Arti Pandey <arti.pandey@dreamtechpress.com>

12.3 Authentication in web applications using the Browser binding 344

JavaScript entry points 344 ▪ Sequence: log in, nextToken, …, log out 344 ▪ Example JavaScript 346

12.4 Authorization and permissions 346

Policies 346 ▪ ACLs 347 ▪ Repository-specific permissions 349 ▪ Changing permissions (applyACL) 349

12.5 Retentions and holds 351

Repository-managed retentions 351 ▪ Client-managed retentions 352 ▪ Holds 352

12.6 Summary 352

13 *Performance 354*

13.1 CMIS performance 354

13.2 Selecting the smallest data set 356

13.3 Performance notes specific to OpenCMIS and DotCMIS 357

13.4 Caching 358

Caching static data 359 ▪ Caching objects 360

13.5 Selecting the fastest binding 364

13.6 Tuning HTTP for CMIS 364

HTTP Keep-Alive 365 ▪ Compression 365 ▪ Authentication and cookies 366 ▪ Timeouts 366

13.7 Summary 366

14 *Building a CMIS server 368*

14.1 Introduction to the OpenCMIS Server Framework 368

CmisService interface 369 ▪ CmisServiceFactory interface 369 The framework 370

14.2 Generating a server stub 370

Building the CMIS server WAR file 372 ▪ Dissecting the CMIS server WAR file 373

14.3 Implementing the CmisServiceFactory interface 374

CmisServiceWrapper 374 ▪ CallContext 375 ▪ Other CmisServiceFactory methods 375

Licensed to Arti Pandey <arti.pandey@dreamtechpress.com>

14.4 Implementing the CmisService interface 376

AbstractCmisService 376 ▪ Best practices for implementing the CmisService 376

14.5 Testing the CMIS server with the OpenCMIS TCK 377

Running the TCK with the CMIS Workbench 378 ▪ TCK results breakdown 379 ▪ Deeper testing 380

14.6 AtomPub differences 380

Providing ObjectInfo 380 ▪ Handling create and delete requests 381 ▪ Dealing with version series 382 ▪ Managing ACLs 382

14.7 Parsing a CMIS query 382

An example of initialization and use 382 ▪ Parsing SELECT 383 ▪ Parsing FROM 384 ▪ Parsing WHERE 384 Parsing ORDER BY 385 ▪ Query wrap-up 385

14.8 Extracting authentication information 386

CallContext 386 ▪ CallContextHandler 386 ▪ Web services 387 ▪ Authentication wrap-up 388

14.9 CMIS extensions 388

14.10 Supporting CMIS 1.0 and CMIS 1.1 389

14.11 Summary 390

appendix A Apache Chemistry OpenCMIS components 391
appendix B BNF 396
appendix C CMIS cheat sheet 410
appendix D Building web applications with JavaScript 412
appendix E References and resources 431
index 433

Licensed to Arti Pandey <arti.pandey@dreamtechpress.com>