The Concordance Database Manual

M. Alan Haley

Apress®

The Concordance Database Manual

Copyright © 2006 by M. Alan Haley

All rights reserved. No part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage or retrieval system, without the prior written permission of the copyright owner and the publisher.

ISBN-13: 978-1-59059-603-6

ISBN-10: 1-59059-603-X

Printed and bound in the United States of America 9 8 7 6 5 4 3 2 1

Trademarked names may appear in this book. Rather than use a trademark symbol with every occurrence of a trademarked name, we use the names only in an editorial fashion and to the benefit of the trademark owner, with no intention of infringement of the trademark.

Lead Editor: Jim Sumser
Technical Reviewer: Sean King
Editorial Board: Steve Anglin, Ewan Buckingham, Gary Cornell, Jason Gilmore, Jonathan Gennick, Jonathan Hassell, James Huddleston, Chris Mills, Matthew Moodie, Dominic Shakeshaft, Jim Sumser, Keir Thomas, Matt Wade
Project Manager: Sofia Marchant
Copy Edit Manager: Nicole LeClerc
Copy Editor: Susannah Pfalzer
Assistant Production Director: Kari Brooks-Copony
Production Editor: Katie Stence
Compositor: Linda Weidemann, Wolf Creek Press
Proofreader: Elizabeth Berry
Indexer: Valerie Perry
Artist: April Milne
Cover Designer: Kurt Krames
Manufacturing Director: Tom Debolski

Distributed to the book trade worldwide by Springer-Verlag New York, Inc., 233 Spring Street, 6th Floor, New York, NY 10013. Phone 1-800-SPRINGER, fax 201-348-4505, e-mail `orders-ny@springer-sbm.com`, or visit `http://www.springeronline.com`.

For information on translations, please contact Apress directly at 2560 Ninth Street, Suite 219, Berkeley, CA 94710. Phone 510-549-5930, fax 510-549-5939, e-mail `info@apress.com`, or visit `http://www.apress.com`.

The information in this book is distributed on an “as is” basis, without warranty. Although every precaution has been taken in the preparation of this work, neither the author(s) nor Apress shall have any liability to any person or entity with respect to any loss or damage caused or alleged to be caused directly or indirectly by the information contained in this work.

The source code for this book is available to readers at `http://www.apress.com` in the Source Code section. You will need to answer questions pertaining to this book in order to successfully download the code.

Contents

About the Author xv
About the Technical Reviewer xvii
Acknowledgments xix
Introduction xxi

CHAPTER 1 Introducing Concordance 1

Types of Data That Can Be Collected 1
Paper 2
Electronic Files 3
E-Mail 6
Transcripts and Depositions 7
Image Data 9
Additional Resources 11
Litigation Support Department 11
Sarbanes-Oxley 11
Professional Organizations 12
Online Resources 12
Summary 13

CHAPTER 2 Using and Installing Concordance 15

What Concordance Does 15
A Closer Look at Concordance Database Structure 17
A Sample Concordance Database 17
Interacting with the Sample Database 18
Searching 21
Concordance Database Limitations 23
Loading Data 24
Coordinating with Vendors 24
Installation and Requirements 25
Hardware Requirements 25
Concordance Server Installation: Step by Step 26
Concordance Workstation Installation: Step by Step 29
Summary 32

CHAPTER 3 Managing Data 33

Data Formats 33
Concordance Data 33
ASCII Text 34
Extended ASCII 34
Electronic Files 36
Using Vendors to Assist with Processing Data 42
Why Is a Vendor Necessary? 42
Vendor Costs 43
Setting Standards 45
Summary 45

CHAPTER 4 Creating and Deploying a Concordance Database 47

Creating a New Concordance Database 47
Loading Delimited Data into Concordance 50
Indexing Data 53
Applying Security 54
Creating an Administrator Account 54
Setting Field Permissions 56
Setting Menu Access Permissions 56
Summary 57

CHAPTER 5 Designing Databases and Defining Field Properties 59

Planning 59
File Naming Conventions 60
Field Naming Conventions 60
Useful Administrative Fields 61
Assessing the Size of a Project 65
Examples of Database Structure 67
Determining Required Roles for Users 69
Creating Concordance Databases 69
Creating Databases from Templates 70
Creating Databases from Scratch 71
Assigning an Authority List to a Specific Field 79
Summary 81

CHAPTER 6 Importing and Exporting Data 83

Importing into Concordance 83
Importing Other Concordance Databases 83
Delimited Text 87
E-Documents 95
Transcripts 103
E-Mail 104
Exporting from Concordance 108
Exporting As a Concordance Database 108
Exporting to a Delimited Text File 109
Database Transcripts 111
Database Structure 112
Summary 112

CHAPTER 7 Administrative Functions 113

Indexing Databases 113
Dictionary and Inverted Text Files 114
Indexing vs. Reindexing 115
Optimizing Indexing 115
Scheduling Indexing Tasks During Times of Nonusage 116
Packing Databases and Dictionary Files 116
Packing a Database 116
Packing the Dictionary Files 118
Zapping a Database 118
Deduplicating Records 118
Selecting Duplication Criteria 119
Original vs. Duplicate Tags 120
Security 120
Managing Security 121
Managing Users and Field-Level Permissions 122
Adding Custom Menu Items 127
Concatenation 128
When Is It Necessary to Concatenate a Database? 129
How Concatenation Works 129
The Concordance Programming Language 131
The Structure of a CPL Program 133
Executing a CPL Program 133
Interacting With Other CPL Programs 139
Summary 139

CHAPTER 8 Using a Concordance Database 141

Opening a Database 141
Browse View 144
Next and Previous Hit Buttons 146
Empties 148
Determining Field Types from Browse View 148
Table View 148
Sorting 149
Table Layout 150
Tallying Fields 151
Split Screen 153
Editing Data 154
Tagging Records 157
Applying Tags 160
Annotations in Browse View 162
Adding and Deleting an Annotation 162
Navigating Through Multiple Annotations 163
Attachments 164
Summary 166

CHAPTER 9 Searching 167

Things to Know About Searching 167
Subjective vs. Objective Data 167
Indexed vs. Nonindexed Data 168
Referencing and Saving Searches 169
The Importance of Training: Computers vs. Humans 170
Viewing Search Results 173
Form Search (Query by Example) 177
Building Searches with Connectors 177
Specifying Fields 178
Entering Search Values 178
Search Then Browse vs. Search Then Table 178
Search Syntax Window 179
Searching Subjective Data 180
Tags 180
Issues 181
Notes 181

Using the Search Window 182
Purpose of the Search Window 183
Scope of Searches 183
Entering Searches 184
Tracking Searches 184
Accessing the Dictionary File 185
Accessing Field Names 185
Fuzzy Searches 186
Using the <Quick Search> Field 187
Overview of the <Quick Search> Field 187
Basic Syntax 188
Relational Searches 189
Combining Keyword Searches with Relational Operators 198
Combining Keyword and Relational Searches with Subjective Data 199
Viewing Search Results 200
Saving Searches As Snapshots and Queries 200
Snapshots 201
Queries 202
Summary 203

CHAPTER 10 Printing 205

Printing the Current Document 205
Printing Sets of Records 208
Fields Tab 208
KWIC Tab 208
Formatting Tab 209
Print Tab 212
Creating Formal Reports 213
Report Writer 213
Annotation Report Wizard 228
Annotation Report Dialog 235
Summary 235

CHAPTER 11 Opticon: Introduction, Overview, and Installation 237

Working with Graphical Images 238
Vector Graphics 238
Raster Graphics 238

Using a Vendor to Create Images 241
Deliverables 241
Workflow 242
Installing Opticon 243
Hardware Requirements 244
Opticon Server Installation: Step by Step 244
Opticon Workstation Installation: Step by Step 248
Summary 251

CHAPTER 12 Using Opticon 253

Setting Opticon As the Default Viewer 253
Opticon's Layout 255
Opening Images 257
Viewing Images 259
View Menu 259
Tools Menu 261
Standard Button Bar 261
Image Button Bar 262
Navigating Through Images 262
Page Menu 263
Image Toolbar 265
Using Redlines 265
Global Preferences 266
Redlines Menu 268
Tools Menu 269
File Menu 269
Redlines Toolbar 270
Searching Redlines 277
The Containing Tab 279
The Advanced Tab 281
Printing Images 281
The Print Tab 281
The Header & Footer Tab 284
The Options Tab 285
The Setup Tab 287
Summary 287

CHAPTER 13 Imagebase Management 289

Using Log Files 289
Log File Structure 289
Examples of Log Files 291
Exporting an Imagebase to a Log File 292
Working with the Imagebase Management Dialog 293
Path 293
Redlines 295
Document Breaks 296
Title Bar 298
Imagebase 299
Edit 300
Register - Load 302
Register - Scan 303
Directory 304
Summary 306

CHAPTER 14 Producing Documents in Opticon 307

Production Numbers 307
Redlines 308
Producing Documents 309
Selecting Records from Concordance 312
Producing Documents with the Production Wizard 313
Production Output 328
Final Steps 331
Modifying Relative File Paths 331
Exporting Concordance Data 332
Summary 332

GLOSSARY 333

INDEX 345