DW 2.0
The Architecture for the Next Generation of Data Warehousing
The Morgan Kaufmann Series in Data Management Systems

Joe Celko's Thinking in Sets
Joe Celko

Business Metadata
Bill Inmon, Bonnie O'Neil and Lowell Fryman

Unleashing Web 2.0
Gottfried Vossen and Stephan Hagemann

Enterprise Knowledge Management
David Loshin

Business Process Change, Second Edition
Paul Harmon

IT Manager's Handbook, Second Edition
Bill Holtznieter and Brian Jaffe

Joe Celko's Puzzles and Answers, Second Edition
Joe Celko

Making Shoes for the Cobbler's Children
Charles Betz

Java Data Mining: Strategy, Standard, and Practice
Mark Hornik, Erik Marcade, and Sunil Venkayala

Joe Celko's Analytics and OLAP in SQL
Joe Celko

Data Preparation for Data Mining Using SAS
Mamdouh Refaat

Querying XML: XQuery, XPath, and SQL/XML in Context
Jim Melton and Stephen Buxton

Data Mining: Concepts and Techniques, Second Edition
Jiawei Han and Micheline Kamber

Database Modeling and Design: Logical Design, Fourth Edition
Toby J. Teorey, Sam S. Lightstone, and Thomas P. Nadeau

Foundations of Multidimensional and Metric Data Structures
Hanan Samet

Joe Celko's SQL for Smarties: Advanced SQL Programming, Third Edition
Joe Celko

Moving Objects Databases
Ralf Hartmut Güting and Markus Schneider

Joe Celko's SQL Programming Style
Joe Celko

Data Mining, Second Edition: Concepts and Techniques
Ian Witten and Eibe Frank

Fuzzy Modeling and Genetic Algorithms for Data Mining and Exploration
Earl Cox

Data Modeling Essentials, Third Edition
Graeme C. Simson and Graham C. Witt

Location-Based Services
Jochen Schiller and Agnes Voisard

Database Modeling with Microsoft® Visio for Enterprise Architects
Terry Halpin, Ken Evans, Patrick Hallock, and Bill Maclean

Designing Data-Intensive Web Applications
Stephano Ceri, Piero Fraternali, Aldo Bongio, Marco Brambilla, Sara Comai, and Maristella Maierav๏a

Mining the Web: Discovering Knowledge from Hypertext Data
Soumen Chakrabarti

Advanced SQL: 1999—Understanding Object-Relational and Other Advanced Features
Jim Melton

Database Tuning: Principles, Experiments, and Troubleshooting Techniques
Dennis Shasha and Philippe Bonnet

SQL: 1999—Understanding Relational Language Components
Jim Melton and Alan R. Simon

Information Visualization in Data Mining and Knowledge Discovery
Edited by Usama Fayyad, Georges G. Grinstein and Andreas Wierse

Transactional Information Systems
Gerhard Weikum and Gottfried Vossen

Spatial Databases
Philippe Rigaux, Michel Scholl, and Agnes Voisard

Information Modeling and Relational Database
Terry Halpin

Component Database Systems
Edited by Klaus R. Dittrich and Andreas Geppert

Managing Reference Data in Enterprise Database
Malcolm Chisholm

Understanding SQL and Java Together
Jim Melton and Andrew Eisenberg

Patrick and Elizabeth O'Neil

The Object Data Standard
Edited by R. G. C. Cattell and Douglas Barry

Data on the Web: From Relations to Semistructured Data and XML
Serge Abiteboul, Peter Buneman, and Dan Suciu

Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations
Ian Witten and Eibe Frank

Joe Celko

Joe Celko's Data and Databases: Concepts in Practice
Joe Celko

Developing Time-Oriented Database Applications in SQL
Richard T. Snodgrass

Web Farming for the Data Warehouse
Richard D. Hackathorn

Management of Heterogeneous and Autonomous Database Systems
Edited by Ahmed Elmagarmid, Marek Rusinkiewicz, and Amit Sheth

Object-Relational DBMSs: Second Edition
Michael Stonebraker and Paul Brown, with Dorothy Moore

A Complete Guide to DB2 Universal Database
Don Chamberlin

Universal Database Management: A Guide to Object-Relational Technology
Cynthia Maro Saracco

Readings in Database Systems, Third Edition
Edited by Michael Stonebraker and Joseph M. Hellerstein

Understanding SQL's Stored Procedures: A Complete Guide to SQL/PSM
Jim Melton

Principles of Multimedia Database Systems
V. S. Subrahmanian

Principles of Database Query Processing for Advanced Applications
Clement T. Yu and WeiYi Meng

Advanced Database Systems
Carlo Zaniolo, Stefano Ceri, Christos Faloutsos, Richard T. Snodgrass, V. S. Subrahmanian, and Roberto Zicari

Principles of Transaction Processing
Philip A. Bernstein and Eric Newcomer

Using the New DB2: IBM's Object-Relational Database System
Don Chamberlin

Distributed Algorithms
Nancy A. Lynch

Active Database Systems: Triggers and Rules for Advanced Database Processing
Edited by Jennifer Widom and Stefano Ceri

Migrating Legacy Systems: Gateways, Interfaces, & the Incremental Approach
Michael L. Brodie and Michael Stonebraker

Atomic Transactions
Nancy Lynch, Michael Merritt, William Weihl, and Alan Fekete

Query Processing for Advanced Database Systems
Edited by Johann Christoph Freytag, David Maier, and Gottfried Vossen

Transaction Processing
Jim Gray and Andreas Reuter

Building an Object-Oriented Database System: The Story of O2
Edited by François Bancilhon, Claude Delobel, and Paris Kanelakis

Database Transaction Models for Advanced Applications
Edited by Ahmed K. Elmagarmid

A Guide to Developing Client/Server SQL Applications
Setrag Khoshafian, Arvola Chan, Anna Wong, and Harry K. T. Wong

Edited by Jim Gray

Camelot and Avalon: A Distributed Transaction Facility
Edited by Jeffrey L. Eppinger, Lily B. Murmert, and Alfred Z. Spector

Readings in Object-Oriented Database Systems
Edited by Stanley B. Zdonik and David Maier

DW 2.0: The Architecture for the Next Generation of Data Warehousing
William H. Inmon, Derek Strauss, and Genia Neuhauss
DW 2.0
The Architecture for the Next Generation of Data Warehousing

W. H. Inmon
Forest Rim Technology

Derek Strauss
Gavroshe

Genia Neushloss
Gavroshe
Dedication for Lynn Inmon, my wife and partner
Contents

Preface ........................................................................................................................................... xvii
Acknowledgments ............................................................................................................................ xx
About the Authors ........................................................................................................................... xxi

CHAPTER 1  A brief history of data warehousing and first-generation data warehouses ......................... 1
  Data base management systems ................................................................................................. 1
  Online applications .................................................................................................................... 2
  Personal computers and 4GL technology .................................................................................. 3
  The spider web environment ....................................................................................................... 4
  Evolution from the business perspective .................................................................................. 5
  The data warehouse environment ............................................................................................... 6
  What is a data warehouse? ......................................................................................................... 7
  Integrating data—a painful experience .................................................................................... 7
  Volumes of data ......................................................................................................................... 8
  A different development approach ........................................................................................... 8
  Evolution to the DW 2.0 environment ....................................................................................... 9
  The business impact of the data warehouse ........................................................................... 11
  Various components of the data warehouse environment ...................................................... 11
    ETL—extract/transform/load .................................................................................................... 12
    ODS—operational data store .................................................................................................. 13
    Data mart ............................................................................................................................... 13
    Exploration warehouse .......................................................................................................... 13
  The evolution of data warehousing from the business perspective ........................................ 14
  Other notions about a data warehouse ...................................................................................... 14
  The active data warehouse ........................................................................................................ 15
  The federated data warehouse approach ................................................................................. 16
  The star schema approach ......................................................................................................... 18
  The data mart data warehouse .................................................................................................. 20
  Building a “real” data warehouse ............................................................................................. 21
  Summary .................................................................................................................................. 22

CHAPTER 2  An introduction to DW 2.0 ............................................................................................ 23
  DW 2.0—a new paradigm .......................................................................................................... 24
  DW 2.0—from the business perspective ..................................................................................... 24
  The life cycle of data .................................................................................................................. 27
  Reasons for the different sectors ............................................................................................... 30
  Metadata ...................................................................................................................................... 31
  Access of data ............................................................................................................................. 33
  Structured data/unstructured data ............................................................................................. 34
Contents

Textual analytics ................................................................. 35
Blather............................................................................... 38
The issue of terminology................................................... 38
Specific text/general text.................................................. 40
Metadata—a major component ......................................... 40
Local metadata ................................................................. 43
A foundation of technology............................................... 45
Changing business requirements ....................................... 47
The flow of data within DW 2.0 ......................................... 48
Volumes of data.................................................................. 50
Useful applications............................................................ 51
DW 2.0 and referential integrity......................................... 52
Reporting in DW 2.0........................................................... 53
Summary........................................................................... 53

CHAPTER 3  DW 2.0 components—about the different sectors .................................................. 55

The Interactive Sector ....................................................... 55
The Integrated Sector........................................................ 62
The Near Line Sector........................................................ 71
The Archival Sector........................................................... 76
Unstructured processing .................................................... 86
From the business perspective ........................................... 90
Summary........................................................................... 92

CHAPTER 4  Metadata in DW 2.0 .................................................. 95

Reusability of data and analysis ........................................ 96
Metadata in DW 2.0............................................................ 96
Active repository/passive repository ................................. 99
The active repository ....................................................... 100
Enterprise metadata.......................................................... 101
Metadata and the system of record .................................... 102
Taxonomy......................................................................... 104
Internal taxonomies/external taxonomies......................... 104
Metadata in the Archival Sector ......................................... 105
Maintaining metadata........................................................ 106
Using metadata—an example ............................................ 106
From the end-user perspective .......................................... 109
Summary........................................................................... 110

CHAPTER 5  Fluidity of the DW 2.0 technology infrastructure .................................................. 111

The technology infrastructure.......................................... 112
Rapid business changes.................................................... 114
Contents

The treadmill of change ................................................................. 114
Getting off the treadmill ............................................................ 115
Reducing the length of time for IT to respond ......................... 115
Semantically temporal, semantically static data ....................... 115
Semantically temporal data ....................................................... 116
Semantically stable data ......................................................... 117
Mixing semantically stable and unstable data ......................... 118
Separating semantically stable and unstable data .................... 118
Mitigating business change ....................................................... 119
Creating snapshots of data ...................................................... 120
A historical record ................................................................. 120
Dividing data ........................................................................... 121
From the end-user perspective .................................................. 121
Summary .................................................................................. 122

CHAPTER 6 Methodology and approach for DW 2.0 ......................... 123

Spiral methodology—a summary of key features ....................... 124
The seven streams approach—an overview .............................. 129
Enterprise reference model stream .......................................... 129
Enterprise knowledge coordination stream ............................. 129
Information factory development stream ................................. 133
Data profiling and mapping stream .......................................... 133
Data correction stream ............................................................ 133
Infrastructure stream ............................................................. 133
Total information quality management stream ....................... 134
Summary .................................................................................. 137

CHAPTER 7 Statistical processing and DW 2.0 ................................. 141

Two types of transactions .......................................................... 141
Using statistical analysis .......................................................... 143
The integrity of the comparison ................................................. 144
Heuristic analysis ................................................................. 145
Freezing data ........................................................................... 146
Exploration processing ............................................................ 146
The frequency of analysis ....................................................... 147
The exploration facility ........................................................... 147
The sources for exploration processing ................................... 149
Refreshing exploration data ..................................................... 149
Project-based data ................................................................. 150
Data marts and the exploration facility .................................... 152
A backflow of data ............................................................... 152
Using exploration data internally ............................................. 155
From the perspective of the business analyst................................. 155
Summary............................................................................................. 156

CHAPTER 8  Data models and DW 2.0...................................................... 157
An intellectual road map ................................................................. 157
The data model and business ......................................................... 157
The scope of integration ................................................................. 158
Making the distinction between granular and summarized data ...... 159
Levels of the data model ............................................................... 159
Data models and the Interactive Sector ........................................ 161
The corporate data model ............................................................ 162
A transformation of models ......................................................... 163
Data models and unstructured data ............................................. 164
From the perspective of the business user ..................................... 166
Summary............................................................................................. 167

CHAPTER 9  Monitoring the DW 2.0 environment............................... 169
Monitoring the DW 2.0 environment ............................................. 169
The transaction monitor ............................................................... 169
Monitoring data quality ............................................................... 170
A data warehouse monitor ........................................................... 171
The transaction monitor—response time ..................................... 171
Peak-period processing ............................................................... 172
The ETL data quality monitor .................................................... 174
The data warehouse monitor ....................................................... 176
Dormant data .................................................................................. 177
From the perspective of the business user ..................................... 178
Summary............................................................................................. 179

CHAPTER 10  DW 2.0 and security.......................................................... 181
Protecting access to data ............................................................. 181
Encryption ......................................................................................... 181
Drawbacks ......................................................................................... 182
The firewall ....................................................................................... 182
Moving data offline ........................................................................ 182
Limiting encryption ........................................................................ 184
A direct dump .................................................................................. 184
The data warehouse monitor ....................................................... 185
Sensing an attack ............................................................................ 185
Security for near line data .......................................................... 187
From the perspective of the business user ..................................... 187
Summary............................................................................................. 188
CHAPTER 11  Time-variant data

All data in DW 2.0—relative to time ............................................................... 191
Time relativity in the Interactive Sector ....................................................... 192
Data relativity elsewhere in DW 2.0 ............................................................ 192
Transactions in the Integrated Sector ......................................................... 193
Discrete data ............................................................................................... 194
Continuous time span data ....................................................................... 194
A sequence of records ............................................................................... 196
Nonoverlapping records .......................................................................... 197
Beginning and ending a sequence of records ............................................ 197
Continuity of data ..................................................................................... 198
Time-collapsed data .................................................................................. 198
Time variance in the Archival Sector ........................................................ 199
From the perspective of the end user ......................................................... 200
Summary .................................................................................................... 200

CHAPTER 12  The flow of data in DW 2.0

The flow of data throughout the architecture ......................................... 203
Entering the Interactive Sector .................................................................. 203
The role of ETL .......................................................................................... 205
Data flow into the Integrated Sector .......................................................... 205
Data flow into the Near Line Sector ........................................................... 207
Data flow into the Archival Sector .............................................................. 209
The falling probability of data access ......................................................... 209
Exception-based flow of data ..................................................................... 210
From the perspective of the business user ................................................ 213
Summary .................................................................................................... 214

CHAPTER 13  ETL processing and DW 2.0

Changing states of data .............................................................................. 215
Where ETL fits .......................................................................................... 215
From application data to corporate data .................................................... 216
ETL in online mode ................................................................................... 216
ETL in batch mode .................................................................................... 217
Source and target ....................................................................................... 218
An ETL mapping ....................................................................................... 219
Changing states—an example ................................................................. 219
More complex transformations ................................................................. 221
ETL and throughput .................................................................................. 222
ETL and metadata ..................................................................................... 223
ETL and an audit trail ............................................................................... 223
ETL and data quality ................................................................. 224
Creating ETL ............................................................................ 224
Code creation or parametrically driven ETL ......................... 225
ETL and rejects ....................................................................... 225
Changed data capture ............................................................. 226
ELT ......................................................................................... 226
From the perspective of the business user ......................... 227
Summary ................................................................................. 228

CHAPTER 14  DW 2.0 and the granularity manager .................. 231
The granularity manager ......................................................... 231
Raising the level of granularity .............................................. 232
Filtering data ........................................................................ 232
The functions of the granularity manager ............................... 234
Home-grown versus third-party granularity managers .......... 236
Parallelizing the granularity manager .................................... 237
Metadata as a by-product ...................................................... 237
From the perspective of the business user ......................... 238
Summary ................................................................................. 238

CHAPTER 15  DW 2.0 and performance ..................................... 239
Good performance—a cornerstone for DW 2.0 .................... 239
Online response time ............................................................. 240
Analytical response time ....................................................... 241
The flow of data .................................................................... 241
Queues ................................................................................ 242
Heuristic processing ............................................................. 243
Analytical productivity and response time ......................... 243
Many facets to performance ................................................. 244
Indexing .............................................................................. 245
Removing dormant data ....................................................... 245
End-user education .............................................................. 246
Monitoring the environment ............................................... 246
Capacity planning ............................................................... 247
Metadata .............................................................................. 249
Batch parallelization ........................................................... 249
Parallelization for transaction processing ........................... 250
Workload management ......................................................... 250
Data marts ........................................................................... 251
Exploration facilities .......................................................... 253
Separation of transactions into classes ................................. 253
Service level agreements ..................................................... 254
Protecting the Interactive Sector ............................................................... 254
Partitioning data ........................................................................................ 255
Choosing the proper hardware ................................................................. 255
Separating farmers and explorers ............................................................. 256
Physically group data together .............................................................. 257
Check automatically generated code ....................................................... 257
From the perspective of the business user ............................................... 258
Summary .................................................................................................... 259

CHAPTER 16 Migration ................................................................................. 261

Houses and cities ..................................................................................... 261
Migration in a perfect world ................................................................. 262
The perfect world almost never happens ............................................. 262
Adding components incrementally ....................................................... 262
Adding the Archival Sector ................................................................. 264
Creating enterprise metadata .............................................................. 265
Building the metadata infrastructure .................................................. 266
“Swallowing” source systems .............................................................. 266
ETL as a shock absorber ................................................................. 267
Migration to the unstructured environment ....................................... 267
From the perspective of the business user ............................................... 269
Summary .....................................................................................................270

CHAPTER 17 Cost justification and DW 2.0 ................................................. 271

Is DW 2.0 worth it? .................................................................................. 271
Macro-level justification ........................................................................ 271
A micro-level cost justification ........................................................... 272
Company B has DW 2.0 ................................................................. 273
Creating new analysis ........................................................................... 273
Executing the steps ............................................................................... 274
So how much does all of this cost? ...................................................... 276
Consider company B ........................................................................... 276
Factoring the cost of DW 2.0 .............................................................. 277
Reality of information ........................................................................... 278
The real economics of DW 2.0 ........................................................... 279
The time value of information .......................................................... 279
The value of integration ....................................................................... 280
Historical information ......................................................................... 280
First-generation DW and DW 2.0—the economics ............................... 281
From the perspective of the business user ............................................... 282
Summary .................................................................................................... 282
Data warehousing has been around for about 2 decades now and has become an essential part of the information technology infrastructure. Data warehousing originally grew in response to the corporate need for information—not data. A data warehouse is a construct that supplies integrated, granular, and historical data to the corporation.

But there is a problem with data warehousing. The problem is that there are many different renditions of what a data warehouse is today. There is the federated data warehouse. There is the active data warehouse. There is the star schema data warehouse. There is the data mart data warehouse. In fact there are about as many renditions of the data warehouse as there are software and hardware vendors.

The problem is that there are many different renditions of what the proper structure of a data warehouse should look like. And each of these renditions is architecturally very different from the others. If you were to enter a room in which a proponent of the federated data warehouse was talking to a proponent of the active data warehouse, you would be hearing the same words, but these words would be meaning very different things. Even though the words were the same, you would not be hearing meaningful communication. When two people from very different contexts are talking, even though they are using the same words, there is no assurance that they are understanding each other.

And thus it is with first-generation data warehousing today.

Into this morass of confusion as to what a data warehouse is or is not comes DW 2.0. DW 2.0 is a definition of the next generation of data warehousing. Unlike the term “data warehouse,” DW 2.0 has a crisp, well-defined meaning. That meaning is identified and defined in this book.

There are many important architectural features of DW 2.0. These architectural features represent an advance in technology and architecture beyond first-generation data warehouses. The following are some of the important features of DW 2.0 discussed in this book:

- The life cycle of data within the data warehouse is recognized. First-generation data warehouses merely placed data on disk storage and called it a warehouse. The truth of the matter is that data—once placed in a data warehouse—has its own life cycle. Once data enters the data warehouse it starts to age. As it ages, the probability of access diminishes. The lessening of the probability of access has profound implications on the technology that is appropriate to the management of the data. Another phenomenon that happens is that as data ages, the volume of data increases. In most cases this increase is dramatic. The task of handling large volumes of data with a decreasing probability of access requires special design considerations lest the cost of the data warehouse become prohibitive and the effective use of the data warehouse becomes impractical.
The data warehouse is most effective when containing both structured and unstructured data. Classical first-generation data warehouses consisted entirely of transaction-oriented structured data. These data warehouses provided a great deal of useful information. But a modern data warehouse should contain both structured and unstructured data. Unstructured data is textual data that appears in medical records, contracts, emails, spreadsheets, and many other documents. There is a wealth of information in unstructured data. But unlocking that information is a real challenge. A detailed description of what is required to create the data warehouse containing both structured and unstructured data is a significant part of DW 2.0.

For a variety of reasons metadata was not considered to be a significant part of first-generation data warehouses. In the definition of second-generation data warehouses, the importance and role of metadata are recognized. In the world of DW 2.0, the issue is not the need for metadata. There is, after all, metadata that exists in DBMS directories, in business objects universes, in ETL tools, and so forth. What is needed is enterprise metadata, where there is a cohesive enterprise view of metadata. All of the many sources of metadata need to be coordinated and placed in an environment where they work together harmoniously. In addition, there is a need for the support of both technical metadata and business metadata in the DW 2.0 environment.

Data warehouses are ultimately built on a foundation of technology. The data warehouse is shaped around a set of business requirements, usually reflecting a data model. Over time the business requirements of the organization change. But the technical foundation underlying the data warehouse does not easily change. And therein lies a problem—the business requirements are constantly changing but the technological foundation is not changing. The stretch between the changing business environment and the static technology environment causes great tension in the organization. In this section of the book, the discussion focuses on two solutions to the dilemma of changing business requirements and static technical foundations of the data warehouse. One solution is software such as Kalido that provides a malleable technology foundation for the data warehouse. The other solution is the design practice of separating static data and temporal data at the point of data base definition. Either of these approaches has the very beneficial effect of allowing the technical foundation of the data warehouse to change while the business requirements are also changing.

There are other important topics addressed in this book. Some of the other topics that are addressed include the following:

- Online update in the DW 2.0 data warehouse infrastructure.
- The ODS. Where does it fit?
- Research processing and statistical analysis against a DW 2.0 data warehouse.
- Archival processing in the DW 2.0 data warehouse environment.
- Near-line processing in the DW 2.0 data warehouse environment.
Data marts and DW 2.0.
■ Granular data and the volumes of data found in the data warehouse.
■ Methodology and development approaches.
■ Data modeling for DW 2.0.

An important feature of the book is the diagram that describes the DW 2.0 environment in its entirety. The diagram—developed through many consulting, seminar, and speaking engagements—represents the different components of the DW 2.0 environment as they are placed together. The diagram is the basic architectural representation of the DW 2.0 environment.

This book is for the business analyst, the information architect, the systems developer, the project manager, the data warehouse technician, the data base administrator, the data modeler, the data administrator, and so forth. It is an introduction to the structure, contents, and issues of the future path of data warehousing.

March 29, 2007
WHI
DS
EN
Acknowledgments

Derek Strauss would like to thank the following family members and friends:

My wife, Denise, and my daughter, Joni, for their understanding and support;
My business partner, Genia Neushloss, and her husband, Jack, for their friendship and support;
Bill Inmon, for his mentorship and the opportunity to collaborate on DW 2.0 and other initiatives;
John Zachman, Dan Meers, Bonnie O’Neil, and Larissa Moss for great working relationships over the years.

Genia Neushloss would like to thank the following family members and friends:

My husband, Jack, for supporting me in all my endeavors;
My sister, Sima, for being my support system;
Derek Strauss, the best business partner anyone could wish for;
Bill Inmon, for his ongoing support and mentorship and the opportunity to collaborate on this book;
John Zachman, Bonnie O’Neil, and Larissa Moss for great working relationships over the years.
About the Authors

W. H. Inmon, the father of data warehousing, has written 49 books translated into nine languages. Bill founded and took public the world’s first ETL software company. He has written over 1000 articles and published in most major trade journals.

Bill has conducted seminars and spoken at conferences on every continent except Antarctica. He holds nine software patents. His latest company is Forest Rim Technology, a company dedicated to the access and integration of unstructured data into the structured world. Bill’s web site—inmoncif.com—attracts over 1,000,000 visitors a month. His weekly newsletter at b-eye-network.com is one of the most widely read in the industry and goes out to 75,000 subscribers each week.

Derek Strauss is founder, CEO, and a principal consultant of Gavroshe. He has 28 years of IT industry experience, 22 years of which were in the information resource management and business intelligence/data warehousing fields.

Derek has initiated and managed numerous enterprise programs and initiatives in the areas of business intelligence, data warehousing, and data quality improvement. Bill Inmon’s Corporate Information Factory and John Zachman’s Enterprise Architecture Framework have been the foundational cornerstones of his work. Derek is also a Specialist Workshop Facilitator. He has spoken at several local and international conferences on data warehousing issues. He is a Certified DW 2.0 Architect and Trainer.

Genia Neushloss is a co-founder and principal consultant of Gavroshe. She has a strong managerial and technical background spanning over 30 years of professional experience in the insurance, finance, manufacturing, mining, and telecommunications industries.

Genia has developed and conducted training courses in JAD/JRP facilitation and systems reengineering. She is a codeveloper of a method set for systems reengineering. She has 22 years of specialization in planning, analyzing, designing, and building data warehouses. Genia has presented before audiences in Europe, the United States, and Africa. She is a Certified DW 2.0 Architect and Trainer.