# Contents

*List of Contributors*  
page vii

   Ian Gorton and Deborah K. Gracio  
   page 1

2. **Anatomy of Data-Intensive Computing Applications**  
   Ian Gorton and Deborah K. Gracio  
   page 12

3. **Hardware Architectures for Data-Intensive Computing Problems: A Case Study for String Matching**  
   Antonino Tumeo, Oreste Villa, and Daniel Chavarría-Miranda  
   page 24

4. **Data Management Architectures**  
   Terence Critchlow, Ghaleb Abdulla, Jacek Becla, Kerstin Kleese-Van Dam, Sam Lang, and Deborah L. McGuinness  
   page 48

5. **Large-Scale Data Management Techniques in Cloud Computing Platforms**  
   Sherif Sakr and Anna Liu  
   page 85

6. **Dimension Reduction for Streaming Data**  
   Chandrika Kamath  
   page 124

7. **Binary Classification with Support Vector Machines**  
   Patrick Nichols, Bobbie-Jo Webb-Robertson, and Christopher Oehmen  
   page 157

8. **Beyond MapReduce: New Requirements for Scalable Data Processing**  
   Bill Howe and Magdalena Balazinska  
   page 180
Contents

9 Let the Data Do the Talking: Hypothesis Discovery from Large-Scale Data Sets in Real Time 235
Christopher Oehmen, Scott Dowson, Wes Hatley, Justin Almquist, Bobbie-Jo Webb-Robertson, Jason McDermott, Ian Gorton, and Lee Ann McCue

10 Data-Intensive Visual Analysis for Cyber-Security 258
William A. Pike, Daniel M. Best, Douglas V. Love, and Shawn J. Bohn

Index 287