

FOREWORD

by David Corrigan

Big data is certainly a hot topic right now. Most organizations are only beginning their adventure with big data. Many are experimenting with new technology, but few are thinking ahead for long-term success. And while most organizations are pursuing big data in order to improve their competitiveness, the fact is fewer still will truly succeed in gaining competitive advantage. The reason is simple: you cannot compete on analytics alone. After all, what does analytics analyze? Information. Information needs to be trusted in order to be acted upon. That is the role of Information Governance: to create trusted information to make big data analytics more successful.

First, there is the technology aspect of Information Governance. IBM® InfoSphere® is the leading platform for creating trusted information. It has deep capabilities in data quality, privacy and security, master data management, information integration, information lifecycle management, metadata management, and policy management. Those capabilities have been well integrated to address multiple requirements. InfoSphere is a key component of IBM's big data platform, and it provides a foundation of trusted information for big data analytics. It is integrated with other big data components for Hadoop-based analytics, streaming analytics, data warehousing, and federated discovery and navigation. The importance of this integration cannot be emphasized enough, as it helps to integrate big data technologies within the enterprise and also to supply them with a steady flow of trusted information. IBM has many clients in multiple industries and has deep industry experience in implementing big data and Information Governance.

Second, there is the process, or strategy, of Information Governance. The importance and the role of a governance strategy are still not well-understood. Information Governance is a business strategy that has a series of IT deliverables. Sunil has been one of the pioneers in this area, defining the Unified Information Governance Process several years ago. He defined several key steps, such as identifying a business problem and executive sponsor, setting up cross-functional governance boards, and measuring and communicating success. He has applied this process at hundreds of clients and has helped them achieve successful implementations. His approach can also be applied to governing big data. It has helped many organizations get the business involved

in governance and establish trusted information for a key enterprise application. In short, this process helps you move beyond an IT project toward a true business strategy. It helps by getting business executives and owners involved in the process of governing data. It helps ensure successful outcomes.

Sunil, thank you for continuing to contribute to the discipline of Information Governance and move it into the new era of computing—the era of big data. And to the readers of this book, remember that the competitive advantage you seek from insights garnered from big data has *two* components: big data analytics and trusted information. Information Governance creates trusted information from very uncertain sources, enabling you to trust and act upon the insights from analytics. I wish you well in your big data strategy.

David Corrigan
Director, Product Marketing, InfoSphere
IBM

FOREWORD

by Inderpal Bhandari

Editor's Note: These remarks were originally published in the book *Big Data Governance: An Emerging Imperative* (Sunil Soares, MC Press, October 2012), from which some of the material in this book has been excerpted.

We now live in an age of seemingly unlimited data. It has slowly, but surely, pervaded our lives. We rely on it to accomplish all manner of tasks, ranging from governing economies and advancing science, to maintaining an electronic record of our health. We have come to realize that its value must be truly understood and unlocked by deriving insights that are revealed through analysis and then translating those insights into information, knowledge, and ultimately action. Recently, with the advent of social media, sensor networks, streaming technology, and the like, the sheer scale of work required to unlock this value has increased beyond the scope of traditional database and data warehousing technologies. In other words, we have entered the age of Big Data.

What is Big Data? What is Big Data Governance? And how does one create business value through Big Data? These and many more questions about Big Data are answered within *Big Data Governance: An Emerging Imperative* (MC Press, October 2012). If you have not yet asked these questions, you *need* to read this book or miss out on the opportunity to differentiate and grow your business by leveraging big data.

Through detailed case studies, Sunil does an excellent job of educating the reader on a complicated and still-evolving subject. For example, he shares a story where a retail outlet uses social media data to analyze consumer awareness and sentiment for a wide range of products. That understanding was used by the retailer to optimize discounts for the products, greatly increasing the firm's profitability.

Sunil shares over twenty such case studies in his book. While reading these case studies, I encourage you to go through a two-step mental exercise, as I did. First, consider how one might attempt to attain the same results as described in the case study, but by using traditional data technologies and processes. Returning to the retail example, by commissioning market research surveys, one can understand the awareness and market sentiment for a range of products, but that process will take months. The research results, when available, are no

longer valid for setting discounts effectively, as awareness and sentiment may be completely different by then. We are now in an era where social media data must be harnessed, using Big Data approaches, to provide business intelligence that is far more timely than is possible through traditional approaches. Second, consider the gaps that exist in current data management processes and systems. Again in the retail example, it is clear that we must be able to match social media data to the products, but the matching rules in traditional master data management systems might not be equal to that task. For instance, these matching rules require natural language processing capability. That is a gap. We then need to provide governance around such rules. Yet another gap. And so on.

Sunil makes this comparative exercise easy by identifying and discussing several such gaps for each case study. The end result is that one comes away with a sound understanding of Big Data Governance, along with answers to the questions I raised earlier in this foreword.

Enough said. Read this book. It will help you enter the age of Big Data.

Inderpal Bhandari
Vice President and Chief Data Officer
Express Scripts