

Wiley CIO Series

Second Edition

INFORMATION GOVERNANCE

**CONCEPTS, STRATEGIES AND
BEST PRACTICES**

Robert F. Smallwood
with leading experts

WILEY

INFORMATION GOVERNANCE

Founded in 1807, John Wiley & Sons is the oldest independent publishing company in the United States. With offices in North America, Europe, Asia, and Australia, Wiley is globally committed to developing and marketing print and electronic products and services for our customers' professional and personal knowledge and understanding.

The Wiley CIO series provides information, tools, and insights to IT executives and managers. The products in this series cover a wide range of topics that supply strategic and implementation guidance on the latest technology trends, leadership, and emerging best practices.

Titles in the Wiley CIO series include:

The Agile Architecture Revolution: How Cloud Computing, REST-Based SOA, and Mobile Computing Are Changing Enterprise IT by Jason Bloomberg

Architecting the Cloud: Design Decisions for Cloud Computing Service Models (SaaS, PaaS, and IaaS) by Michael Kavis

Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses by Michael Minelli, Michele Chambers, and Ambiga Dhiraj

The Chief Information Officer's Body of Knowledge: People, Process, and Technology by Dean Lane

Cloud Computing and Electronic Discovery by James P. Martin and Harry Cendrowski

Confessions of a Successful CIO: How the Best CIOs Tackle Their Toughest Business Challenges by Dan Roberts and Brian Watson

CIO Best Practices: Enabling Strategic Value with Information Technology (Second Edition) by Joe Stenzel, Randy Betancourt, Gary Cokins, Alyssa Farrell, Bill Flemming, Michael H. Hugos, Jonathan Hujsak, and Karl Schubert

The CIO Playbook: Strategies and Best Practices for IT Leaders to Deliver Value by Nicholas R. Colisto

Decoding the IT Value Problem: An Executive Guide for Achieving Optimal ROI on Critical IT Investments by Gregory J. Fell

Enterprise Performance Management Done Right: An Operating System for Your Organization by Ron Dimon

Information Governance: Concepts, Strategies and Best Practices by Robert F. Smallwood

IT Leadership Manual: Roadmap to Becoming a Trusted Business Partner by Alan R. Guibord

Leading the Epic Revolution: How CIOs Drive Innovation and Create Value Across the Enterprise by Hunter Muller

Managing Electronic Records: Methods, Best Practices, and Technologies by Robert F. Smallwood

On Top of the Cloud: How CIOs Leverage New Technologies to Drive Change and Build Value Across the Enterprise by Hunter Muller

Straight to the Top: CIO Leadership in a Mobile, Social, and Cloud-based World (Second Edition) by Gregory S. Smith

Strategic IT: Best Practices for Managers and Executives by Arthur M. Langer and Lyle Yorks

Trust and Partnership: Strategic IT Management for Turbulent Times by Robert Benson, Piet Ribbers, and Ronald Billstein

Transforming IT Culture: How to Use Social Intelligence, Human Factors, and Collaboration to Create an IT Department That Outperforms by Frank Wander

Unleashing the Power of IT: Bringing People, Business, and Technology Together, Second Edition by Dan Roberts

The U.S. Technology Skills Gap: What Every Technology Executive Must Know to Save America's Future by Gary J. Beach

INFORMATION GOVERNANCE

CONCEPTS, STRATEGIES, AND
BEST PRACTICES

SECOND EDITION

Robert F. Smallwood

WILEY

Copyright © 2020 by Robert F. Smallwood. All rights reserved.

Chapter 7 © 2014 by Barclay Blair.

Portions of Chapter 8 © 2014 by Randolph Kahn.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.

Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 646-8600, or on the Web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at www.wiley.com/go/permissions.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993, or fax (317) 572-4002.

Wiley publishes in a variety of print and electronic formats and by print-on-demand. Some material included with standard print versions of this book may not be included in e-books or in print-on-demand. If this book refers to media such as a CD or DVD that is not included in the version you purchased, you may download this material at <http://booksupport.wiley.com>. For more information about Wiley products, visit www.wiley.com.

Library of Congress Cataloging-in-Publication Data:

Names: Smallwood, Robert F., 1959- author.

Title: Information governance: concepts, strategies, and best practices /
Robert F. Smallwood.

Description: Second edition. | Hoboken, New Jersey: John Wiley & Sons, Inc.,
[2020] | Series: The Wiley CIO series | Includes index. |

Identifiers: LCCN 2019015574 (print) | LCCN 2019017654 (ebook) | ISBN
9781119491415 (Adobe PDF) | ISBN 9781119491408 (ePub) | ISBN 9781119491446
(hardback)

Subjects: LCSH: Information technology—Management. | Management information
systems. | Electronic records—Management.

Classification: LCC HD30.2 (ebook) | LCC HD30.2 .S617 2020 (print) | DDC
658.4/038—dc23

LC record available at <https://lccn.loc.gov/2019015574>

Cover Design: Wiley

Cover Image: ©style_TTT/Shutterstock

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

*For my sons
and the next generation of tech-savvy managers*

CONTENTS

PREFACE XVII

ACKNOWLEDGMENTS XIX

PART ONE—Information Governance Concepts, Definitions, and Principles **1**

CHAPTER 1 The Information Governance Imperative 3

- Early Development of IG 4
- Big Data Impact 5
- Defining Information Governance 7
- IG Is Not a Project, But an Ongoing Program 9
- Why IG Is Good Business 9
- Failures in Information Governance 11
- Form IG Policies, Then Apply Technology for Enforcement 14

CHAPTER 2 Information Governance, IT Governance, Data Governance: What's the Difference? 19

- Data Governance 19
- Data Governance Strategy Tips 20
- IT Governance 21
- IT Governance Frameworks 22
- Information Governance 25
- Impact of a Successful IG Program 25
- Summing Up the Differences 26

CHAPTER 3 Information Governance Principles 29

- The Sedona Conference® Commentary on Information Governance 29
- Smallwood IG Principles 30
- Accountability Is Key 34
- Generally Accepted Recordkeeping Principles® 35
Contributed by Charmaine Brooks
- Assessment and Improvement Roadmap 42
- Information Security Principles 45
- Privacy Principles 45
- Who Should Determine IG Policies? 48

Part Two—Information Governance Risk Assessment and Strategic Planning **53**

CHAPTER 4 Information Asset Risk Planning and Management 55

- The Information Risk Planning Process 56
- Create a Risk Profile 59
- Information Risk Planning and Management Summary 65

CHAPTER 5 Strategic Planning and Best Practices for Information Governance 69

- Crucial Executive Sponsor Role 70
- Evolving Role of the Executive Sponsor 71
- Building Your IG Team 72
- Assigning IG Team Roles and Responsibilities 72
- Align Your IG Plan with Organizational Strategic Plans 73
- Survey and Evaluate External Factors 75
- Formulating the IG Strategic Plan 81

CHAPTER 6 Information Governance Policy Development 87

- The Sedona Conference IG Principles 87
- A Brief Review of Generally Accepted Recordkeeping Principles® 88
- IG Reference Model 88
- Best Practices Considerations 91
- Standards Considerations 92
- Benefits and Risks of Standards 93
- Key Standards Relevant to IG Efforts 93
- Major National and Regional ERM Standards 98
- Making Your Best Practices and Standards Selections to Inform Your IG Framework 105
- Roles and Responsibilities 105
- Program Communications and Training 106
- Program Controls, Monitoring, Auditing, and Enforcement 107

Part Three—Information Governance Key Impact Areas **113**

CHAPTER 7 Information Governance for Business Units 115

- Start with Business Objective Alignment 115
- Which Business Units Are the Best Candidates to Pilot an IG Program? 117
- What Is Infonomics? 117
- How to Begin an IG Program 118
- Business Considerations for an IG Program 119
- By Barclay T. Blair*

Changing Information Environment	119
Calculating Information Costs	121
Big Data Opportunities and Challenges	122
Full Cost Accounting for Information	123
Calculating the Cost of Owning Unstructured Information	124
The Path to Information Value	127
Challenging the Culture	129
New Information Models	129
Future State: What Will the IG-Enabled Organization Look Like?	130
Moving Forward	132

CHAPTER 8 Information Governance and Legal Functions 135

Robert Smallwood with Randy Kahn, Esq., and Barry Murphy

Introduction to E-Discovery: The Revised 2006 and 2015 Federal Rules of Civil Procedure Changed Everything	135
Big Data Impact	137
More Details on the Revised FRCP Rules	138
Landmark E-Discovery Case: <i>Zubulake v. UBS Warburg</i>	139
E-Discovery Techniques	140
E-Discovery Reference Model	140
The Intersection of IG and E-Discovery	143
<i>By Barry Murphy</i>	
Building on Legal Hold Programs to Launch Defensible Disposition	146
<i>By Barry Murphy</i>	
Destructive Retention of E-Mail	147
Newer Technologies That Can Assist in E-Discovery	147
Defensible Disposal: The Only Real Way to Manage Terabytes and Petabytes	151
<i>By Randy Kahn, Esq.</i>	

CHAPTER 9 Information Governance and Records and Information Management Functions 161

Records Management Business Rationale	163
Why Is Records Management So Challenging?	165
Benefits of Electronic Records Management	166
Additional Intangible Benefits	167
Inventorizing E-Records	168
RM Intersection with Data Privacy Management	169
<i>By Teresa Schoch</i>	
Generally Accepted Recordkeeping Principles®	171
E-Records Inventory Challenges	172

Records Inventory Purposes	172
Records Inventorying Steps	173
Appraising the Value of Records	184
Ensuring Adoption and Compliance of RM Policy	184
Sample Information Asset Survey Questions	190
General Principles of a Retention Scheduling	191
Developing a Records Retention Schedule	192
Why Are Retention Schedules Needed?	193
What Records Do You Have to Schedule? Inventory and Classification	195
Rationale for Records Groupings	196
Records Series Identification and Classification	197
Retention of E-Mail Records	197
How Long Should You Keep Old E-Mails?	199
Destructive Retention of E-Mail	199
Legal Requirements and Compliance Research	200
Event-Based Retention Scheduling for Disposition of E-Records	201
Prerequisites for Event-Based Disposition	202
Final Disposition and Closure Criteria	203
Retaining Transitory Records	204
Implementation of the Retention Schedule and Disposal of Records	204
Ongoing Maintenance of the Retention Schedule	205
Audit to Manage Compliance with the Retention Schedule	206

CHAPTER 10 Information Governance and Information Technology Functions 211

Data Governance	213
Steps to Governing Data Effectively	214
Data Governance Framework	215
Information Management	216
IT Governance	220
IG Best Practices for Database Security and Compliance	223
Tying It All Together	225

CHAPTER 11 Information Governance and Privacy and Security Functions 229

Information Privacy	229
<i>By Andrew Ysasi</i>	
Generally Accepted Privacy Principles	231
Fair Information Practices (FIPS)	232
OCED Privacy Principles	233
Madrid Resolution 2009	234

EU General Data Protection Regulation	235
GDPR: A Look at Its First Year	237
<i>By Mark Driskill</i>	
Privacy Programs	239
Privacy in the United States	240
Privacy Laws	244
Cybersecurity	245
Cyberattacks Proliferate	246
Insider Threat: Malicious or Not	247
Information Security Assessments and Awareness Training	248
<i>By Baird Brueseke</i>	
Cybersecurity Considerations and Approaches	253
<i>By Robert Smallwood</i>	
Defense in Depth	254
Controlling Access Using Identity Access Management	254
Enforcing IG: Protect Files with Rules and Permissions	255
Challenge of Securing Confidential E-Documents	256
Apply Better Technology for Better Enforcement in the Extended Enterprise	257
E-Mail Encryption	259
Secure Communications Using Record-Free E-Mail	260
Digital Signatures	261
Document Encryption	262
Data Loss Prevention (DLP) Technology	262
Missing Piece: Information Rights Management (IRM)	265
Embedded Protection	268
Hybrid Approach: Combining DLP and IRM Technologies	270
Securing Trade Secrets After Layoffs and Terminations	270
Persistently Protecting Blueprints and CAD Documents	271
Securing Internal Price Lists	272
Approaches for Securing Data Once It Leaves the Organization	272
Document Labeling	274
Document Analytics	275
Confidential Stream Messaging	275

Part Four—Information Governance for Delivery Platforms **283**

CHAPTER 12 Information Governance for E-Mail and Instant Messaging **285**

Employees Regularly Expose Organizations to E-Mail Risk	286
E-Mail Polices Should Be Realistic and Technology Agnostic	287

E-Record Retention: Fundamentally a Legal Issue	287
Preserve E-Mail Integrity and Admissibility with Automatic Archiving	288
Instant Messaging	291
Best Practices for Business IM Use	292
Technology to Monitor IM	293
Tips for Safer IM	294
Team and Channel Messaging Solutions Emerge	294

CHAPTER 13 Information Governance for Social Media 299

Dr. Patricia Franks and Robert Smallwood

Types of Social Media in Web 2.0	299
Additional Social Media Categories	303
Social Media in the Enterprise	304
Key Ways Social Media Is Different from E-Mail and Instant Messaging	305
Biggest Risks of Social Media	306
Legal Risks of Social Media Posts	307
Tools to Archive Social Media	309
IG Considerations for Social Media	311
Key Social Media Policy Guidelines	312
Records Management and Litigation Considerations for Social Media	313
Emerging Best Practices for Managing Social Media Records	315

CHAPTER 14 Information Governance for Mobile Devices 319

Current Trends in Mobile Computing	322
Security Risks of Mobile Computing	323
Securing Mobile Data	324
Mobile Device Management (MDM)	324
IG for Mobile Computing	325
Building Security into Mobile Applications	326
Best Practices to Secure Mobile Applications	330
Developing Mobile Device Policies	330

CHAPTER 15 Information Governance for Cloud Computing 335

Monica Crocker and Robert Smallwood

Defining Cloud Computing	336
Key Characteristics of Cloud Computing	337
What Cloud Computing Really Means	338
Cloud Deployment Models	339
Benefits of the Cloud	340
Security Threats with Cloud Computing	341

Managing Documents and Records in the Cloud	351
IG Guidelines for Cloud Computing Solutions	351
IG for SharePoint and Office365	352
<i>By Robert Bogue</i>	

CHAPTER 16 Leveraging and Governing Emerging Technologies 357

Data Analytics	357
Descriptive Analytics	358
Diagnostic Analytics	358
Predictive Analytics	358
Prescriptive Analytics	359
Which Type of Analytics Is Best?	359
Artificial Intelligence	363
The Role of Artificial Intelligence in IG	363
Blockchain: A New Approach with Clear Advantages	366
<i>By Darra Hoffman</i>	
Breaking Down the Definition of Blockchain	366
The Internet of Things: IG Challenges	372
IoT as a System of Contracts	375
IoT Basic Risks and IG Issues	376
IoT E-Discovery Issues	377
Why IoT Trustworthiness Is a Journey and <i>Not</i> a Project	380
<i>By Bassam Zarkout</i>	
Governing the IoT Data	381
IoT Trustworthiness	382
Information Governance Versus IoT Trustworthiness	384
IoT Trustworthiness Journey	385
Conclusion	386

Part Five—Long-Term Program Issues 391

CHAPTER 17 Long-Term Digital Preservation 393

<i>Charles M. Dollar and Lori J. Ashley</i>	
Defining Long-Term Digital Preservation	393
Key Factors in Long-Term Digital Preservation	394
Threats to Preserving Records	396
Digital Preservation Standards	397
PREMIS Preservation Metadata Standard	404
Recommended Open Standard Technology–Neutral Formats	405
Digital Preservation Requirements	409
Long-Term Digital Preservation Capability Maturity Model®	409

Scope of the Capability Maturity Model	412
Digital Preservation Capability Performance Metrics	416
Digital Preservation Strategies and Techniques	417
Evolving Marketplace	419
Looking Forward	420
Conclusion	421

CHAPTER 18 Maintaining an Information Governance Program and Culture of Compliance 425

Monitoring and Accountability	425
Change Management—Required	426
<i>By Monica Crocker</i>	
Continuous Process Improvement	429
Why Continuous Improvement Is Needed	430

APPENDIX A Information Organization and Classification: Taxonomies and Metadata 433

Barb Blackburn, CRM, with Robert Smallwood; edited by Seth Earley

Importance of Navigation and Classification	435
When Is a New Taxonomy Needed?	435
Taxonomies Improve Search Results	436
Metadata and Taxonomy	437
Metadata Governance, Standards, and Strategies	438
Types of Metadata	440
Core Metadata Issues	441
International Metadata Standards and Guidance	442
Records Grouping Rationale	446
Business Classification Scheme, File Plans, and Taxonomy	446
Classification and Taxonomy	447
Prebuilt Versus Custom Taxonomies	448
Thesaurus Use in Taxonomies	449
Taxonomy Types	449
Business Process Analysis	453
Taxonomy Testing: A Necessary Step	457
Taxonomy Maintenance	457
Social Tagging and Folksonomies	458

APPENDIX B Laws and Major Regulations Related to Records Management 463

United States	463
Gramm-Leach-Bliley Act	463

Healthcare Insurance Portability and Accountability Act of 1996 (HIPAA)	463
PATRIOT Act (Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism Act of 2001)	464
Sarbanes-Oxley Act (SOX)	464
SEC Rule 17A-4	464
CFR Title 47, Part 42—Telecommunications	464
CFR Title 21, Part 11—Pharmaceuticals	464
US Federal Authority on Archives and Records: National Archives and Records Administration (NARA)	465
US Code of Federal Regulations	465
Canada	466
United Kingdom	468
Australia	469
Identifying Records Management Requirements in Other Legislation	471
APPENDIX C Laws and Major Regulations Related to Privacy	475
United States	475
European Union General Data Protection Regulation (GDPR)	476
Major Privacy Laws Worldwide, by Country	478
GLOSSARY	481
ABOUT THE AUTHOR	499
ABOUT THE MAJOR CONTRIBUTORS	501
INDEX	505

PREFACE

In the five plus years since the first edition of this book was published, information governance (IG) has matured as a discipline, and business executives and managers at leading enterprises now see IG programs as increasingly valuable. A combination of factors have created an imperative for IG programs: new, tightened regulations; the continuing deluge of Big Data; and the realization that new value can be gained from information stores using analytics have all combined to raise the profile of IG programs across the globe.

In particular, new privacy legislation, including the EU General Data Protection Regulation and the California Consumer Privacy Act, helped foster a newfound awareness of data protection issues, and organizations worldwide scrambled to inventory and gain insight into their information stores. This is often a first step in IG programs, and so the realization of IG as a needed and valued undertaking set in. Enterprises began to see IG not only as a cost center and risk reduction activity, but also as one that can add value to the enterprise, in some cases even monetizing information.

This book clarifies and codifies what IG is—and what it is not—and how to launch, control, and manage IG programs. Based on exhaustive research, and with the contributions of a number of industry pioneers and experts, this book lays out IG as a complete discipline, fully updated, including an expanded section on information privacy and new material on managing emerging technologies.

IG is a “super-discipline” of sorts in that it includes components of privacy, cybersecurity, infonomics, law and e-discovery, records management, compliance, risk management, information technology (IT), business operations, and more. This unique blend calls for a new breed of information professional who is competent across these complex disciplines. Training and education are key to IG program success, and this book provides the fundamentals as well as advanced concepts to enable organizations to train a new generation of IG professionals. The book is being used to guide IG programs at major corporations, as well as to educate graduate students in information science, computer science, law, and business.

Practitioners in the component areas of IG will find the book useful in expanding their knowledge and helping them understand the linkages between the various facets of IG. And how breaking down existing siloed approaches and leveraging information as an asset across the enterprise is critical to gaining the full benefits of IG programs.

The book strives to offer clear and concise IG concepts, actionable strategies, and proven best practices in an understandable and digestible way; a concerted effort was made to simplify language and offer examples. There are summaries of key points throughout the book and at the end of each chapter to help the reader retain key points. The text is organized into five parts: (1) IG Concepts, Definitions, and Principles; (2) IG Risk Assessment and Strategic Planning; (3) IG Key Impact Areas; (4) IG for Information Delivery Platforms, including a new section on emerging technologies; and (5) Long-Term Program Issues.

No other book offers comprehensive coverage of the complex and challenging field of IG with such clarity. Use the insights and advice contained in these pages and your IG program will have lower risks and costs, and produce better and more measurable results.

Robert Smallwood

ACKNOWLEDGMENTS

I would like to gratefully thank my colleagues for the support and generous contributions of their expertise and time, which made this updated and comprehensive text possible.

Many thanks to Lori Ashley, Jason R. Baron, Barb Blackburn, Barclay Blair, Robert Bogue, Charmaine Brooks, Baird Brueseke, Ken Chasse, Monica Crocker, Charles Dollar, Mark Driskill, Seth Early, Sam Fossett, Dr. Patricia Franks, Randy Kahn, Dennis Kessler, Darra Hoffman, Doug Laney, Paula Lederman, Reynold Leming, Barry Murphy, Robert Seiner, Teresa Schoch, Andrew Ysasi, and Bassam Zarkout.

I am truly honored to include their insightful work and owe them a great debt of gratitude.

PART ONE

Information Governance Concepts, Definitions, and Principles

CHAPTER 1

The Information Governance Imperative

Effective **information governance** (IG) programs improve operational efficiency and compliance capabilities while leveraging information as an asset to maximize their value. Active IG programs are the hallmark of well-managed organizations, and increasingly IG has become an imperative, especially for global enterprises.

A “perfect storm” of compliance pressures, cybersecurity concerns, Big Data volumes, and the increasing recognition that information itself has value have contributed to a substantial increase in the number of organizations implementing IG programs.

Most significantly, the European Union (EU) General Data Protection Regulation (GDPR), which went into effect May 25, 2018, left companies across the globe scrambling to gain control over the consumer data they had housed. The GDPR legislation applies to all citizens in the EU and the European Economic Area (EEA), regardless of where they reside, and also visitors and temporary residents of the EU. So any global enterprise doing business with EU/EEA citizens—or even visitors—must comply with the legislation or face a major fine. The primary goal of GDPR is to give citizens control over their personal data.

Brought about in part because of GDPR compliance concerns, membership in the International Association of Privacy Professionals (IAPP) grew from around 25,000 members in 2017 to over 40,000 members in 2018, and it continues to grow.

A first step in the GDPR compliance process is to conduct an inventory of an enterprise’s information assets to create a data map showing where all incidences of data are housed. This is commonly the first major implementation step in IG programs, so the IG discipline and support for IG programs made substantial strides in 2018 with the lead-up to GDPR going into effect. Then California passed its California Consumer Privacy Act (CCPA), which borrowed many concepts from GDPR and required that any company (of a certain size) handling the personally identifiable information (PII) of California residents (in specified volumes) comply by January 1, 2020. Suddenly US-based companies could no longer ignore privacy regulations, and the momentum for IG programs that could manage privacy compliance requirements accelerated.

During this same time frame, data breaches and ransomware attacks became more prevalent and damaging, and organizations adopted IG programs to reduce the likelihood of cyberattacks, and their impact. IG programs implement effective risk reduction countermeasures.

A first step in the GDPR compliance process is to conduct an inventory of an enterprise's information assets to create a data map.

Added to that has been the continued massive increase on overall data volumes that organizations must manage, which results in managing a lot of unknown “dark data,” which lacks metadata and has not been classified. Organizations also retain large volumes of redundant, outdated, and trivial (ROT) information that needs to be identified and disposed of. Cleaning up the ROT that organizations manage reduces their overall storage footprint and costs, and makes information easier to find, leading to improved productivity for knowledge workers.

IG programs are also about optimizing and finding new value in information. The concept of managing and monetizing information is core to the emerging field of **infonomics**, which is the discipline that assigns “economic significance” to information and provides a framework to manage, measure, and monetize information.¹ Gartner's former analyst Doug Laney published a groundbreaking book in 2018, *Infonomics*, which delineates infonomics principles in great detail, providing many examples of ways organizations have harvested new value by finding ways to monetize information or leverage its value.

Infonomics is the discipline that assigns “economic significance” to information and provides a framework to manage, measure, and monetize information.

Early Development of IG

IG has its roots in the United Kingdom's healthcare system. Across the pond, the government-funded National Health Service (NHS) has focused on IG to ensure data quality and protect patient data since 2002. Although IG was mentioned in journals and scholarly articles decades ago, the UK is arguably the home of healthcare IG, and perhaps the IG discipline.² Could this be the reason the UK leads the world in healthcare quality? Certainly, it must be a major contributing factor.

The United States has the most expensive healthcare in the world, the most sophisticated equipment, the most advanced medicines, the best-trained doctors—yet in a recent study of healthcare quality, the United States came in dead last out of 11 civilized nations.³ The UK, Switzerland, and Sweden topped the list.

The U.S. healthcare problem is not due to poor training, inferior equipment, inferior medicines, or lack of financial resources. No, the problem is likely primarily *a failure to get the right information to the right people at the right time*; that is, caregivers must have accurate, current clinical information to do their jobs properly. These are IG issues.

Since 2002 each UK healthcare organization has been tasked with completing the IG Toolkit, managed by NHS Digital for the UK Department of Health. Although the IG Toolkit has evolved over the years, its core has remained constant. However, in

April 2018 it was replaced with a new tool, the Data Security and Protection Toolkit, based around 10 National Data Security Standards that have been formulated by the UK's National Data Guardian.⁴

Big Data Impact

According to the research group Gartner, Inc., Big Data is defined as “. . . high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making.”⁵ A practical definition should also include the idea that the amount of information—both **structured data** (in databases) and **unstructured information** (e.g. e-mail, scanned documents, PDFs, MS Office documents) is so massive that it cannot be processed using traditional database tools (e.g. DB2, SQL) and analytic software techniques.⁶

In today's information overload era of Big Data—characterized by massive growth in business data volumes and velocity—the ability to distill key insights from enormous amounts of data is a major business differentiator and source of sustainable competitive advantage. In fact, a report by the World Economic Forum stated that data is a new asset class and personal data is “the new oil.”⁷ And we are generating more than we can manage effectively with current methods and tools.

The Big Data numbers are overwhelming: Estimates and projections vary, but it has been stated that 90% of the data existing worldwide today was created in the past two years,⁸ and that every two days more information is generated than was from the dawn of civilization until 2003.⁹ This trend will continue.

Certainly, there are new and emerging opportunities arising from the accumulation and analysis of all that data we are busy generating and collecting. New enterprises are springing up to capitalize on data mining and business analytics opportunities. Back in 2012, the US federal government joined in, announcing \$200 million in Big Data research programs.¹⁰

The onslaught of Big Data necessitates that IG be implemented to discard unneeded data in a legally defensible way.

However, established organizations, especially larger ones, are being crushed by this onslaught of Big Data: it is just too expensive to keep all the information that is being generated, and unneeded and ROT information becomes a sort of irrelevant sludge of data debris for decision makers to wade through. They have difficulty knowing which information is accurate and meaningful “signal,” and which is simply irrelevant “noise.” This means they do not have the precise information on which they can base good business decisions.

And it has real costs: the burden of massive stores of information has increased storage costs dramatically, caused overloaded systems to fail, and increased legal discovery costs.¹¹ Furthermore, the longer that data is kept the more likely that it will need to be migrated to newer computing platforms, driving up conversion costs; and

legally, there is the risk that somewhere in that mountain of data an organization keeps is a piece of information that represents a significant legal liability.¹²

This is where the worlds of Big Data and business collide. For Big Data proponents, more data is always better, and there is no perceived downside to the accumulation of massive amounts of data. In the business world, though, the realities of legal **e-discovery** mean the opposite is true.¹³ To reduce risk, liability, and costs, it is critical for unneeded or useless information to be disposed of in a systematic, methodical, and “legally defensible” (justifiable in legal proceedings) way, when it no longer has legal, regulatory, or business value.

Big Data values massive accumulation of data whereas in business, e-discovery realities and potential legal liabilities dictate that data be culled down to only that which has clear business value.

Organizations are struggling to reduce and right-size their information footprint by discarding superfluous and redundant data, **e-documents**, and information. *But the critical issue is devising policies, methods, and processes, and then deploying information technology (IT) to sort through the information and determine what is valuable and what no longer has value and can be discarded.*

IT, compliance, and legal representatives in organizations have a clear sense that most of the information stored is unneeded, raises costs, and poses risks. According to a survey by the Compliance, Governance and Oversight Council (CGOC), respondents estimated that approximately one-quarter of information stored in organizations has real business value, while 5% must be kept as business records, and about 1% is retained due to a litigation hold.¹⁴ This means that [about] 69% of information in most companies has no business, legal or regulatory value. “Companies that are able to dispose of this debris return more profit to shareholders, can use more of their IT budgets for strategic investments, and can avoid excess expense in legal and regulatory response” [italics added].

Only about one-quarter of the information that organizations are managing has real business value.

*With a smaller **information footprint**, organizations can more easily find what they need and derive business value from it.*¹⁵ They must eliminate the data debris regularly and consistently, and to do this, processes and systems must be in place to cull out valuable information and discard the data debris. An IG program sets the framework to accomplish this.

The business environment has also underscored the need for IG. According to Ted Friedman at Gartner, “The recent global financial crisis has put information governance in the spotlight... [it] is a priority of IT and business leaders as a result of various pressures, including regulatory compliance mandates and the urgent need for improved decision-making.”¹⁶

And IG mastery is critical for executives: many CIOs in regulated industries have been fired from their jobs for failed IG initiatives.¹⁷

With a smaller information footprint, it is easier for organizations to find the information they need and derive business value from it.

Defining Information Governance

Information governance is a sort of “super discipline” that has emerged as a result of new and tightened legislation governing businesses, privacy concerns, legal demands, external pressures such as hacking and data breaches, and the recognition that multiple overlapping disciplines were needed to address today’s information management challenges in an increasingly regulated and litigated business environment.¹⁸

IG is a subset of corporate governance, and includes key concepts from information security, data privacy and protection, records and information management (RIM), content management, IT and data governance, risk management, litigation readiness, regulatory compliance, **long-term digital preservation (LTDP)**, and even analytics and information economics, (infonomics). This also means that it includes related technology and discipline subcategories such as **document management**, enterprise search, knowledge management, and **disaster recovery (DR)/business continuity (BC)**.

Information governance is a subset of corporate governance.

Practicing good IG is the essential foundation for building legally defensible disposition practices to discard unneeded information, and to secure confidential, sensitive, and secret information, which may include trade secrets, strategic plans, price lists, blueprints, or personal information subject to privacy laws. Good IG provides the basis for consistent, reliable methods for managing, securing, controlling, and optimizing information.

Having trusted and reliable records, reports, data, and databases allows managers to make key decisions with confidence.¹⁹ And accessing that information and data analytics insights in a timely fashion can yield a long-term sustainable competitive advantage, creating more agile enterprises.

IG is a sort of “super discipline” that encompasses a variety of key concepts from a variety of related disciplines.

To do this, organizations must standardize and systematize their handling of information, and audit their processes to ensure so. They must analyze and optimize how information is accessed, controlled, managed, shared, stored, preserved, and audited. They must have complete, current, and relevant policies, processes, and technologies to manage and control information, including *who* is able to access what information, and *when*, to meet external legal and regulatory demands and internal governance policy requirements. The idea is to provide the right information to the right people at the right time—securely. *Security, control, and optimization of information*; this, in short, is IG.

Practicing good IG is the essential foundation for building legally defensible disposition practices to discard unneeded information.

Information governance is a subset of corporate governance, which has been around as long as corporations have existed. IG is a rather new multidisciplinary field that is still being defined, but has gained significant traction in the past several years. The focus on IG comes not only from privacy, cybersecurity, compliance, legal, and records management functionaries, but also from executives who understand they are accountable for the governance of information, and that theft or erosion of information assets has real costs and consequences. It can cause corporate brand equity to collapse, and stock price to tumble.

IG is an all-encompassing term for *how an organization manages the totality of its information*.

Information governance programs are about minimizing information risks and costs, while maximizing its value. In short, IG is the security, control, and optimization of information.

Information governance programs are about minimizing information risks and costs, while maximizing its value. IG is control of information to meet business, legal, regulatory, and risk demands.

Stated differently, information governance is “a quality-control discipline for managing, using, improving, and protecting information.”²⁰