

# Contents

Figures	xi
Tables	xiii
Foreword by Michael Lesk	xv
Preface	xix
Acknowledgments	xxiii
<b>Part I: Introduction to Digital Preservation</b>	
1 What Is Digital Preservation?	3
Digital Preservation Is <i>Not</i> . . .	3
Digital Preservation Is Not Only about Backups and Recovery	3
Digital Preservation Is Not Only about Access	4
Digital Preservation Is Not an Afterthought	5
Elements of Digital Preservation	6
Why Digital Preservation?	6
Digital Preservation: A Management Issue	10
Why Libraries, Archives, Museums?	12
Conclusion	13
2 Getting Started with the Digital Preservation Triad	17
Steps in the Digital Preservation Process	19
The Digital Preservation Triad	21
Management	21
Policies and Planning for Digital Preservation	21
Technology Decisions	23

The Question of Rights	24
Resource Issues	25
Outreach and Sustainability	29
Technology	30
Trustworthy Digital Preservation Systems	30
Metadata	33
File Formats	34
Content	34
Copyright Issues	35
Content-Related Challenges	36
Conclusion	37
<b>Part II: Management Aspects</b>	
3 The OAIS Reference Model	43
History	44
OAIS Reference Model Components	44
Vocabulary	45
Information Model	45
OAIS Functional Model	46
OAIS Required Responsibilities	48
Conclusion	52
4 Human Resources and Education	55
Human Resources	55
Categories of Human Resources	56
Education for Digital Preservation	57
Digital Preservation and Digital Curation: What's in a Name?	58
University-Level Education for Digital Preservation	58
Continuing Education for Digital Preservation	61
Research in Digital Preservation	62
Conclusion	65
5 Sustainable Digital Preservation	67
Digital Preservation as Risk Management	68
Involvement in the Creation Process	69
Open and/or Well-Documented Standards and Systems	69
Documentation of Decisions	69
Accepted Standards for Metadata Schemas	69
Needs of the User	70
Exit Strategy	70
Succession Planning	71
Other Considerations for Risk Management	71
Sustainable Digital Resources	72

Blue Ribbon Task Force on Sustainable Digital Preservation and Access	73
Five Conditions Necessary for Digital Preservation Sustainability	74
Factors Affecting Digital Preservation Sustainability	75
Organizational Factors	76
Financial Factors	77
Social and Societal Factors	80
Technological Factors	81
Homegrown, Open Source, and Proprietary Software	
Development Models	82
Memorandums of Understanding (MOUs)	84
Conclusion	89
<b>Part III: Technology Aspects</b>	
6 The Digital Preservation Repository and Trust	95
Trust	96
Trusted Repository Criteria and Checklists	97
European Framework for Audit and Certification of	
Digital Repositories	98
TRAC, TRD, and ISO 16363	103
DRAMBORA	106
Conclusion	107
7 Metadata and Metadata for Digital Preservation	111
Metadata in Digital Librarianship	112
Descriptive Metadata	114
Administrative Metadata	114
Technical Metadata	114
Structural Metadata	115
Markup Languages	115
Structure of Metadata Files	117
Metadata Schema	118
Application Profiles	122
Converting Records and Data to a New Format	122
Metadata Generation and Creation	123
Documentation	125
Metadata Necessary for Digital Preservation	125
Preservation Description Information (PDI)	127
Digital Preservation Metadata	131
Metadata Specific to Digital Preservation	132
PREMIS Model	133
Metadata Encoding and Transmission Standard (METS)	136
METS Profiles	137
Conclusion	137

8	File Formats and Software for Digital Preservation	143
	File Formats	144
	File Formats for Digital Preservation	145
	Evaluating File Formats for Digital Preservation	148
	Determining File Formats	155
	File Extensions	155
	MIME Internet Media Types	156
	File format Registries	156
	Why Are Registries So Difficult?	159
	Software to Help Identify File Formats	159
	Generic Tools	159
	File Type Specific Tools	162
	Conclusion	164
<b>Part IV: Content-Related Aspects</b>		
9	Collection Development	171
	Criteria	173
	Existing Collections	173
	New Collections	173
	Conclusion	175
10	Preserving Research Data	179
	Research Data	180
	Research Data Life Cycle	180
	Big Data	183
	Small Data as Big Data's Counterpart	185
	Metadata Schema for Science Data	185
	Directory Interchange Format (DIF)	185
	The Content Standard for Digital Geospatial Metadata (CSDGM)	186
	Darwin Core Schema	186
	Core Scientific Metadata Model	186
	Harvestable Scientific Metadata	187
	Open Data Initiatives	187
	The U.S. National Science Foundation	188
	The U.S. National Institute of Health	189
	Other U.S. Initiatives	189
	English-Speaking Countries: Approaches to Open Data	190
	Human Subjects and Data Preservation	191
	Challenges with Preserving Human Subject Data	191
	Conclusion	192
11	Preserving Humanities Content	197
	Computerizing the Humanities	199
	Big Data in the Digital Humanities	199

Funding for the Digital Humanities	200
Humanities Sources	200
Metadata Schema for Published Texts	201
Metadata Schema for Digital Texts	202
Metadata Schema for Encoding Visual Resources: Museum Artifacts	203
Metadata Schema for Encoding Video and Sound	205
Conclusion	206
12 Conclusion	211
Appendix A Select Resources in Support of Digital Preservation	213
Selected Digital Preservation Organizations	213
Selected Digital Preservation Consortium/Group Initiatives	214
Data Preservation	214
Other Initiatives	215
Reports	216
General Reports on Digital Preservation	216
Archives	217
Museums	217
Metadata	217
File Formats	217
Moving Images	218
Music	218
Webliographies and Webinars	218
Webliographies	218
Webinars	219
Books, Guides, and Textbooks	219
Online Digital Preservation Glossaries	220
Directories for Digital Preservation Education	220
Centers Supporting Research and Teaching in Digital Preservation	221
Conferences and In-Person Events	221
Core Conferences on Digital Preservation	221
Related Conferences on Digital Preservation	222
Glossary	223
Bibliography	235
Index	259