

**DIGITAL ELECTRONIC ARCHIVING:  
THE STATE OF THE ART AND  
THE STATE OF THE PRACTICE**

**A Report Sponsored by:**

**International Council for Scientific and Technical Information  
Information Policy Committee**

**And**

**CENDI**

**Prepared by:**

**Gail Hodge and Bonnie C. Carroll  
Information International Associates, Inc.  
Oak Ridge, TN**

**Final**

**April 26, 1999**

## TABLE OF CONTENTS

EXECUTIVE SUMMARY .....	1
Introduction .....	15
What is Digital Electronic Archiving? .....	17
Purpose and Scope of the Study .....	17
Data Collection and Analysis .....	18
Methodology .....	18
Highlighted Projects .....	20
General State of the Art/Practice .....	24
Identified Organizational Models .....	25
Data Centers .....	26
The Centralized Data Center .....	26
Federated Data Centers .....	27
Cooperative Data Centers .....	30
The Institutional Archive .....	31
Third Party Repositories .....	33
Repository Management Agents .....	33
Publication Service Providers .....	36
Legal Depositories .....	38
National Libraries .....	38
National Archives .....	39
Interoperable Archives: Open Archival Information System Model .....	42
Life Cycle Players and Their Roles .....	43
Creators/Producers .....	43
Publishers .....	44
Secondary Services .....	45
Libraries and Library Consortia .....	46
Funding Agencies .....	48
Users .....	48
Best Practices by Life Cycle Function .....	48
Creation .....	48
Acquisition and Collection Development .....	49
Determining Extent .....	51

Archiving Related Links .....	52
Cataloging and Identification .....	53
Metadata .....	53
Ensuring Persistence through Identification .....	54
Storage .....	55
Hardware/Software Migration .....	55
Refreshing the Media .....	57
Backup and Recovery .....	57
Preservation .....	57
Refreshing the Site Contents .....	57
Retention .....	58
Standards, Transformations vs. Native Formats .....	58
Preserving the Look and Feel .....	59
Access .....	60
Access Mechanisms .....	60
Rights Management .....	61
Practices Related to Specific Formats and Data Types .....	61
Character Sets for Numeric and Textual Information .....	62
Resolution and Compression Considerations for Images .....	62
Object Archiving for Multimedia .....	63
Practices Related to Specific Object (Document) Types .....	64
Biological Sequence Data .....	65
Documentation for Software and Datasets .....	65
Cost/Resources .....	66
Conclusions .....	67
Policies .....	67
Organizational Models .....	68
Economic Models .....	70
Multiple Models in a Networked Environment .....	70
Recommended Next Steps .....	71
ICSTI .....	71
Both ICSTI and CENDI .....	73
Bibliography .....	76
APPENDIX A-1: Initial Survey .....	A-1
APPENDIX A-2: Follow-up Discussion Questions .....	A-5
APPENDIX B: General Description of All Projects .....	B-1

APPENDIX C: Detailed Description of Selected Projects ..... C-1

APPENDIX D-1: Contributors List ..... D-1

APPENDIX D-2: Additional Contacts List ..... D-3