16 AXIOMS OF DIGITAL PRESERVATION

Some of the axioms relating to a digital repository

1. A REPOSITORY IS NOT A PIECE OF SOFTWARE

Software cannot preserve anything. Software cannot be a repository in itself. A repository is the sum of financial resources, hardware, staff time, and ongoing implementation of policies and planning to ensure long-term access to content. Any software system you use to enable you preserving and providing access to digital content is by necessity temporary. You need to be able to get your stuff out of it because it likely will not last forever. Similarly, there is no software that “does” digital preservation.


Rubens Chiri / Divulgação Osesp

2. INSTITUTIONS MAKE PRESERVATION POSSIBLE

Each of us will die. Without care and management, the things that mattered to us will persist for some period of time related to the durability of their mediums. With that noted, the primary enablers of preservation for the long term are our institutions (libraries, archives, museums, families, religious organizations, governments, etc.) As such, the possibility of preservation is enabled through the design and function of those institutions. Their org charts, hiring practices, funding, credibility, etc. are all key parts of the cultural machinery that makes preservation possible.

Shutterstock Images

3. TOOLS CAN GET IN THE WAY JUST AS MUCH AS THEY CAN HELP

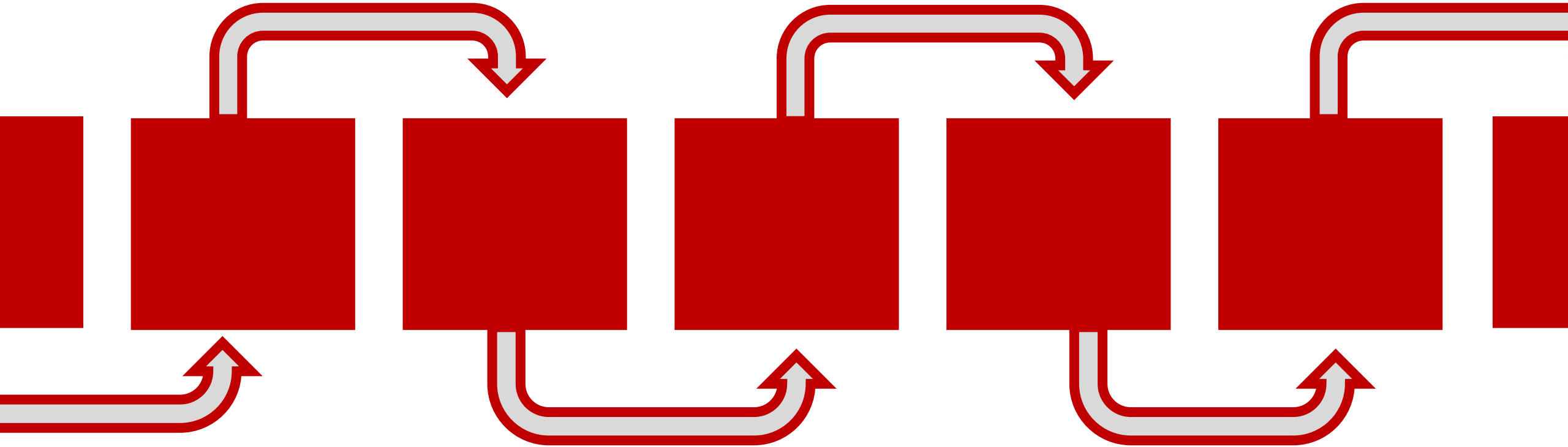
Specialized digital preservation tools and software are just as likely to get in the way of solving your digital preservation problems as they are to help. In many cases, it's much more straightforward to start small and implement simple and discrete tools and practices to keep track of your digital information using nothing more than the file system you happen to be working in. It's better to start simple and then introduce tools that help you improve your process then to simply buy into some complex system without having gotten your house in order first.

(SOME) ORGANIZATIONS ADDRESSING DIGITAL PRESERVATION STANDARDS AND TOOLS

Digital POWRR Tool Evaluation Grid	Ingest					Processing						Access		Storage					Maintenance			Other		
	Copy	Fixity Check	Virus Scan	File Dedupe	Auto Unique ID	Auto Metadata Creation	Auto Metadata Harvest	Manual Metadata	Rights Management	Package Metadata	Auto SIP Creation	Public Interface	Auto DIP Creation	Auto AIP Creation	Robotic Long Term Bit Preservation	Redundancy	Geographic Redundancy	Exit Strategy	Migration	Monitoring	Auto Recovery	Open Source	Clear Documentation	Cost
**Internet Archive	x	x			x	x	x	x	x		x	x	x	x	x	x	x					x		Free
iText																						x		Free
JHOVE (JSTOR/Harvard Object Validation Environment)																						x	x	Free
JHOVE 2						x	x															x	x	Free
Jpylyzer																						x	x	Free
MarcAlizer																						x		Free
**MetaArchive (A private LOCKSS Network)	x	x		x	x			x		x	x			x	x	x	x	x		x	x		x	Varies
Metadata Extraction Tool						x	x	x		x		x										x	x	Free
MIXED (Migration to Intermediate XML for Electronic Data)																			x					Free
NARA File Analyzer and Metadata Harvesting Tool						x	x															x	x	Free
KEEP Emulation Framework																						x	x	Free

4. NOTHING *HAS BEEN* PRESERVED, THERE ARE ONLY THINGS *BEING* PRESERVED.

Preservation is the result of ongoing work of people and commitments of resources. The work is never finished. This has significant ramifications for how we think about staffing and resourcing preservation work. If you want to evaluate how serious an organization is about digital preservation don't start by looking at their code, their storage architecture, or talking to their developers. Start by talking to their finance people.



6. BACKING UP DATA IS NOT DIGITAL PRESERVATION.

If you start talking about digital preservation and someone tells you “oh, don’t worry about it, we back everything up nightly” you need to be prepared to explain how and why that is not digital preservation. It’s important to develop your explanation of what the differences are. Many of the aspects that go into backing up data for current use are similar to aspects of digital preservation work but the near term concerns of being able to restore data are significantly different from the long term issues related to ensuring access to content in the future.

Backup is not Digital Preservation

But

Digital Preservation cannot exist without a Backup

9. DIGITAL PRESERVATION IS MAKING THE BEST USE OF RESOURCES TO MITIGATE THREATS AND RISKS.

You are never done with digital preservation. It is not something that can be accomplished or finished. Digital preservation is a continual process of understanding the risks you face for losing content or losing the ability to render and interact with it and making use of whatever resources you have to mitigate those risks.

ASANA

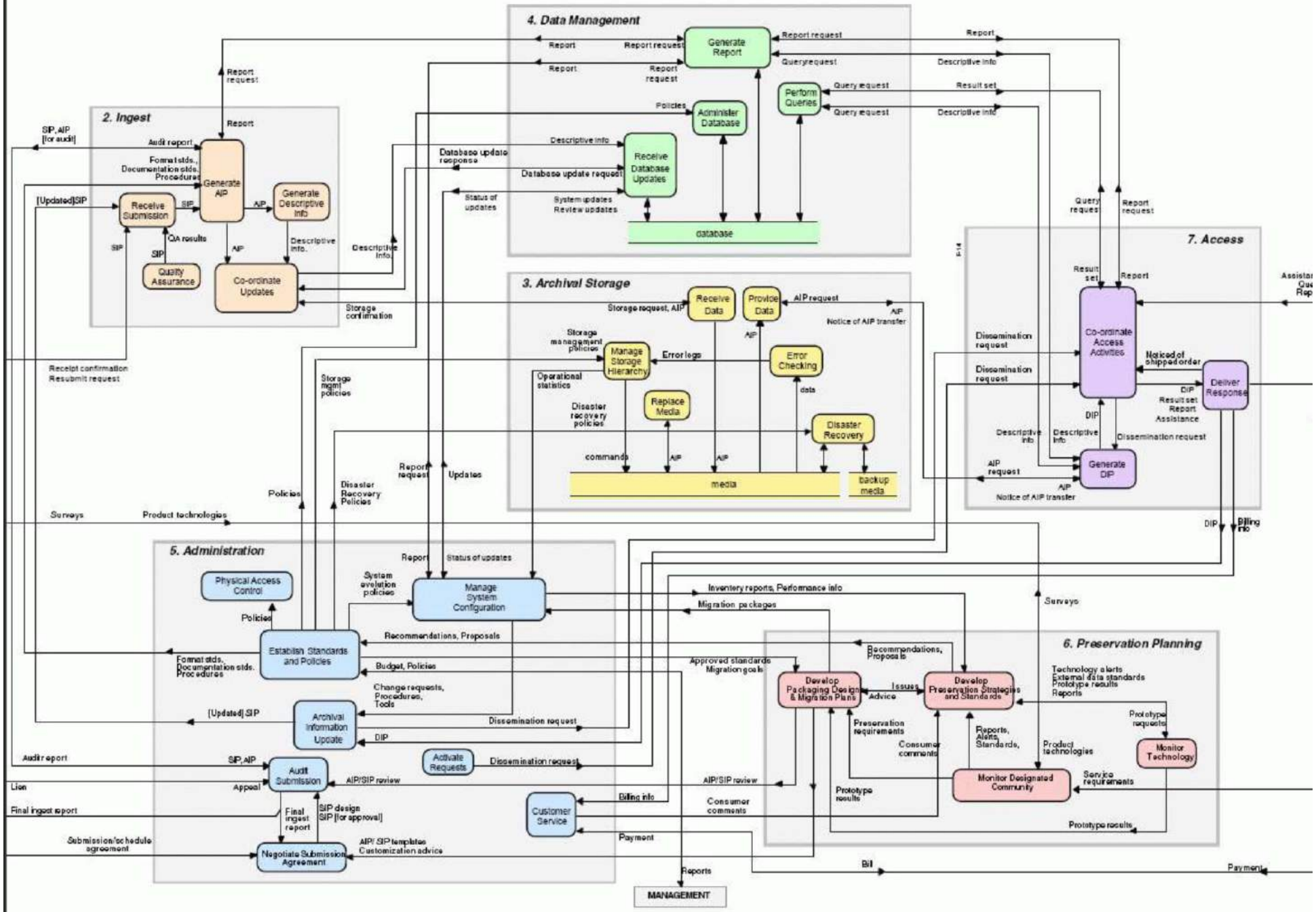
11. IT'S LONG PAST TIME TO START TAKING ACTIONS.

You can read and ponder complicated data models, schemas for tracking and logging preservation actions, and a range of other complex and interesting topics for years but it's not going to help "get the boxes off the floor." There are practical and pragmatic things everyone can and should do now to mitigate many of the most pressing risks of loss. So be sure to prioritize doing those things first before delving into many of the more open-ended areas of digital preservation work and research.

Complex X Simple

12. HIGHLY TECHNICAL DEFINITIONS OF DIGITAL PRESERVATION ARE COMPLICIT IN SILENCING THE PAST

Much of the language and specifications of digital preservation have developed into complex sets of requirements that obfuscate many of the practical things anyone and any organization can do to increase the likelihood of access to content in the future. As such, a highly technical framing of digital preservation has resulted in many smaller and less resource rich institutions feeling like they just can't do digital preservation, or that they need to hire consultants to tell them about complex preservation metadata standards when what they need to do first is make a copy of their files.


4. Data Management
Generate Report
Report request
Report
Query request
Descriptive info
Perform Queries
Result set
Administer Database
Policies
Receive Database Updates
Database update response
Database update request
Status of updates
System updates
Review updates
database
2. Ingest
SIP, AIP [for audit]
Audit report
Format stds., Documentation stds., Procedures
Generate AIP
Generate Descriptive Info
[Updated]SIP
Receive Submission
SIP
AIP
QA results
Quality Assurance
Descriptive info.
Co-ordinate Updates
Storage confirmation
3. Archival Storage
Receive Data
Provide Data
AIP request
Notice of AIP transfer
Storage request, AIP
Storage management policies
Manage Storage Hierarchy
Error logs
Error Checking
Operational statistics
Replace Media
Disaster recovery policies
Disaster Recovery
commands
media
backup media
7. Access
Result set
Co-ordinate Access Activities
Dissemination request
Notice of shipped order
Deliver Response
DIP
Result set
Report
Assistance
Descriptive info
Generate DIP
AIP request
Notice of AIP transfer
Receipt confirmation
Resubmit request
Storage mgmt policies
Disaster Recovery Policies
Surveys
Product technologies
Billing info
5. Administration
Physical Access Control
System evolution policies
Manage System Configuration
Status of updates
Establish Standards and Policies
Recommendations, Proposals
Format stds., Documentation stds. Procedures
Budget, Policies
Change requests, Procedures, Tools
Archival Information Update
[Updated] SIP
Activate Requests
Audit report
SIP, AIP
Audit Submission
AIP/SIP review
Appeal
Lien
Final ingest report
Final ingest report
SIP design
SIP [for approval]
Submission/schedule agreement
Negotiate Submission Agreement
AIP/ SIP templates
Customization advice
Customer Service
Billing info
Payment
Reports
MANAGEMENT
6. Preservation Planning
Develop Packaging Design & Migration Plans
Issues
Advice
Develop Preservation Strategies and Standards
Technology alerts
External data standards
Prototype results
Reports
Approved standards
Migration goals
Inventory reports, Performance info
Migration packages
Preservation requirements
Consumer comments
Reports, Alerts, Standards,
Prototype requests
Monitor Technology
Product technologies
Monitor Designated Community
Service requirements
Prototype results
Bill
Payment

14. ACCEPT AND EMBRACE THE ARCHIVAL SLIVER.

We've never saved everything. We've never saved most things. When we start from the understanding that most things are temporary and likely to be lost to history, we can shift to focus our energy on making sure we line up the resources necessary to protect the things that matter the most. The ideology of "the digital" makes it seem like we could or should attempt to save everything. However, this comes from the mistaken thinking that digital preservation is primarily a technical challenge instead of a social and ethical one.

WHAT TO START TODAY?

SIX THINGS TO DO STARTING TODAY

1. **Identify what digital stuff you've got that you need to keep**
2. **Get the digital boxes off the floor**
3. **Schedule out a plan for improving things and checking in**
4. **Read the NDSA levels of Digital Preservation Paper**
5. **Join the communities of practice**
6. **Read the Theory and Craft of Digital Preservation**