No one tool is the solution to your digital preservation problems

Digital preservation is technology-enabled and human-driven
OAIS and other models inform workflows

workflows combine tools with human action

tools perform functions and populate workflows
Workflows: proto-archivematica

[Diagram showing OAIS: AD-02 Generate AIP]

https://wiki.archivematica.org/OAIS_Activity_Diagrams
Workflows: COW

http://coptr.digipres.org/Workflow:Community_Owned_Workflows
Workflows: OSSArcFlow

https://educopia.org/research/ossarcflow
Your own workflows will be documentary & aspirational.....

aka “As-Is” & “To-Be”
share your workflow experience

(doesn’t have to be digital preservation)
Integrating tools into your workflows

- **Start small.** Basic tools to accomplish individual tasks are an improvement, but only a step away from your manual processes.
- **Refine your tools.** Automate and/or use bundled tools/systems wherever possible and reasonable.
- **Update your workflows.** Define sequence of tool application and actions.
- **Refine your workflows.** String tools together.
Step 1: Build Your Requirements

Once you’ve recorded your As-Is workflows, you can begin identifying tasks that you would like to accomplish with a tool or set of tools.

---Where do you want more automation? More metadata? Where are you happy with current tools? What are your pain points?

ex. An archivist might use `dd` in the command line to copy disk images in a current workflow. In order to make the workflow more inclusive and gather more metadata, they might want to move to something like Forensics Toolkit to accomplish imaging tasks in a friendly interface.
Step 2: Resources for Information about Tools

- Community Owned digital Preservation Tool (COPTR) Registry [http://coptr.digipres.org/Main_Page](http://coptr.digipres.org/Main_Page)
- Preserving (Digital) Objects with Restricted Resources (POWRR) [http://digitalpowrr.niu.edu/](http://digitalpowrr.niu.edu/)
Step 3: Select Tools

What can my organization manage realistically?
(policies, procedures, mandate, support, etc)

What technology could we operate and support?
(which tools fit best with existing infrastructure? can IT help?)

How many of our resources can we dedicate to this work?
(humans, $$$$$$, time)
Step 3: Select Tools

- Interoperability
  - Will it work with other tools you have or will add?

- Right-sized
  - Is there a bunch of extra functionality that you don’t need?
    - Look for the appropriate granularity for problem and setting
  - Is there enough useful documentation and training for you effectively implement the tool?

- Maintenance
  - Can you maintain the tool over inevitable changes in staff, roles?
  - How often does the tool update/change? Will you have the bandwidth to keep up with new or changed functions?
Open Source Mythology

- Free
- Only used by small or under-resourced institutions
- Security is more of a problem than in proprietary tools
- You can’t get support with open source tools
- Not compatible with proprietary tools
Ask for the *same things* from open source and proprietary tools:

openness, transparency, interoperability, security, and support
Individual tools exist for every step in the digital preservation process. We’ll focus on commonly used tools with bundled functionality.
Tools for figuring out what you have

BitCurator

FORENSIC TOOLKIT (FTK)
The Big 4 In Digital Preservation Processing

archivematica

ExLibris Rosetta

libsafe

Preservica
Web Archiving Tools

- ARCHIVE-IT
- Social Feed Manager
- HTTrack WEBSITE COPIER
- Webrecorder
Tools for email

- Emailchemy: http://www.weirdkid.com/products/emailchemy/
- EPADD: https://library.stanford.edu/projects/epadd/about
Preservation Storage
Preservation Storage Criteria

- The Preservation Storage Criteria provides a list of criteria, or design attributes, for storage that supports the work of digital preservation.
  - Version 3 open for comment: [https://osf.io/hztr5/](https://osf.io/hztr5/)
what tools do you use?
what tools do you have questions about?
Common tool challenges

- Installation glitches and other bugs
- Development delays
- Insufficient documentation
- Unresponsive support
- Task and data redundancy - overlapping functions
- Incompatibility
- Learning curve / training requirements
- Data dependence - only renderable/understandable using the tool
- Organizational / IT support (access challenges, root and/or admin permissions)
- Multiple dependencies
Tools for self assessment

- Core Trust Seal
  - [https://www.coretrustseal.org/](https://www.coretrustseal.org/)

- Trustworthy Repositories Audit & Certification (TRAC)
    - download drupal tool from Artefactual Systems wiki: [https://www.archivematica.org/wiki/Internal_audit_tool](https://www.archivematica.org/wiki/Internal_audit_tool)
    - Superseded in 2012 by the ISO 16363 TDR Checklist (Magenta Book)
Pay attention to what the community is doing and share what you’re doing. Engage & participate.
What’s on the horizon

- Machine learning (ex. Predictive Coding at U of Illinois [https://osf.io/6yp9j/](https://osf.io/6yp9j/))
- Ethical challenges to processes and tools (ex. [https://www.docnow.io/](https://www.docnow.io/))
- Private & Sensitive Data Storage
- Environmental sustainability
- One to watch - [https://projectelectron.rockarch.org/](https://projectelectron.rockarch.org/)
questions?

Thanks to the DPM workshops and the AVP/DPN workflow curriculum for content which was adapted for parts of this session.