Towards Automated Digital Preservation

Jack O’Sullivan
OPFCON
9 June 2020
What is Automated Preservation?
Community Knowledge
Four steps towards auto-preservation

**Step 1**
Share best practice with the whole DP community

**Step 2**
Create “Rule Sets” of preservation actions to cover a whole set of formats

E.g. “Convert all high-res original images to a low-res format for access”

**Step 3**
Applying these rule sets to new and existing content at scale

Make it so

**Step 4**
Improving the foundations

More tools for validation, property extraction, render
Step 1: The Preservation Acton Registries API
Preservation Action Registries

http://parcore.org
Sharing best practice for format preservation

What are we sharing

• Tools e.g. ImageMagick (open source image editing tool)
• Actions e.g. ImageMagick can be run with these parameters to produce downsized JPEGs
• Business rule e.g. use this action on TIFF files

Open API

• Currently Preservica, Artefactual, Arkivum, OPF, JISC with more soon
• So we can see other people’s techniques and we can see theirs
• Full documentation online
  https://preview.preservica.com/Registry/par/documentation.html
Step 2: Using PAR info to build recommended rule sets
Building Recommendations

Using Published Policies/Advice

- NARA’s Digital Preservation Framework “Still Images”
- Map a format (e.g. Kodak PhotoCD) to a Business Rule to create TIFF

Longer Term

- Institutions publish these rule sets directly
Step 3:
Applying the rules automatically
Applying the rules

**Rule sets always updated**
- The rule sets are looked up live from the registry so always up to date
- Always applied to new ingests

**Applying to existing data**
- When telling the system that the rules have been updated you can say what you want to happen
  - e.g. Re-migrate specific formats
Auto-preservation in action: migration
What Next?

• Challenges of scale:
  – How does this work with petabytes of data?
  – How does this work with hundreds of users?

• Network challenges:
  – Can we get the expertise published?
Thank you

info@preservica.com
preservica.com
@preservica
@jacko_os