Automation tools in AV workflows

Building digital pipelines using standard IT tools.

Jonáš Svatoš
Národní filmový archiv, CZ

No Time To Wait! #3
London, UK. 25.-26.10.2018

Various steps when working with AV content

- Ingest
- Verification
- Storage
- Transcoding
- Packaging
- Delivery

- ..

Workflow management?

- DAMS
- Digipres systems like Archivematica, Preservica, Adlib
- Manual flow

Scripts are good, but..

- Individual scripts do not serve a generic purpose
- Complex workflows must chain many steps together
- Logging anyone?
- Auditing the logs anyone?

Is this deterministic?

Národní filmový archiv

```
(venv) [digilab@dl1-render-1 ~]$ cd /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV/
(venv) [digilab@dl1-render-1 ZBOROV]$ ll
total 204140708
-rwxr--r--. 1 1028 users   963133204 22. říj 16.09 Zborov_20181022_original_scan.wav
-rw-r--r--. 1 1024 users 68012932294 16. říj 18.29 ZBOROV_ALL_REELS_GRADED_WIP_181015_1080_24p_LT.mov
-rw-r--r--. 1 1024 users 68969349901 17. říj 20.52 ZBOROV_ALL_REELS_GRADED_WIP_181016_1080_24p_LT.mov
-rw-r--r--. 1 1024 users 69167624548 23. říj 13.08 ZBOROV_bez_orezu_1080_24p_422LT_MONOfinal_181023.mov
-rwxr--r--. 1 1028 users   961585926 22. říj 16.10 Zborov_restoration_20181022_bez_sumu_na_zacatku.wav
-rw-rw-r--. 1 1024 users   965447656 16. říj 10.12 Zborov_restoration_WIP_20181011.wav
(venv) [digilab@dl1-render-1 ZBOROV]$ bagit.py --md5 --sha1 .
2018-10-20 06:34:25,971 - INFO - Creating bag for directory /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV
2018-10-20 06:34:25,975 - INFO - Creating data directory
2018-10-20 06:34:25,977 - INFO - Moving Zborov_20181022_original_scan.wav to /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV/tmpnm4jw228/Zborov_20181022_original_scan.wav
2018-10-20 06:34:25,977 - INFO - Moving ZBOROV_ALL_REELS_GRADED_WIP_181015_1080_24p_LT.mov to /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV/tmpnm4jw228/ZBOROV_ALL_REELS_GRADED_WIP_181015_1080_24p_LT.mov
2018-10-20 06:34:25,978 - INFO - Moving ZBOROV_ALL_REELS_GRADED_WIP_181016_1080_24p_LT.mov to /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV/tmpnm4jw228/ZBOROV_ALL_REELS_GRADED_WIP_181016_1080_24p_LT.mov
2018-10-20 06:34:25,980 - INFO - Moving ZBOROV_bez_orezu_1080_24p_422LT_MONOfinal_181023.mov to /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV/tmpnm4jw228/ZBOROV_bez_orezu_1080_24p_422LT_MONOfinal_181023.mov
2018-10-20 06:34:25,981 - INFO - Moving Zborov_restoration_20181022_bez_sumu_na_zacatku.wav to /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV/tmpnm4jw228/Zborov_restoration_20181022_bez_sumu_na_zacatku.wav
2018-10-20 06:34:25,982 - INFO - Moving Zborov_restoration_WIP_20181011.wav to /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV/tmpnm4jw228/Zborov_restoration_WIP_20181011.wav
2018-10-20 06:34:25,983 - INFO - Moving /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV/tmpnm4jw228 to data
2018-10-20 06:34:25,984 - INFO - Using 1 processes to generate manifests: md5, sha1
2018-10-20 06:34:25,985 - INFO - Generating manifest lines for file data/ZBOROV_ALL_REELS_GRADED_WIP_181015_1080_24p_LT.mov
2018-10-20 06:38:25,313 - INFO - Generating manifest lines for file data/ZBOROV_ALL_REELS_GRADED_WIP_181016_1080_24p_LT.mov
2018-10-20 06:42:28,468 - INFO - Generating manifest lines for file data/ZBOROV_bez_orezu_1080_24p_422LT_MONOfinal_181023.mov
2018-10-20 06:46:31,476 - INFO - Generating manifest lines for file data/Zborov_20181022_original_scan.wav
2018-10-20 06:46:34,917 - INFO - Generating manifest lines for file data/Zborov_restoration_20181022_bez_sumu_na_zacatku.wav
2018-10-20 06:46:38,264 - INFO - Generating manifest lines for file data/Zborov_restoration_WIP_20181011.wav
2018-10-20 06:46:41,848 - INFO - Creating bagit.txt
2018-10-20 06:46:41,896 - INFO - Creating bag-info.txt
2018-10-20 06:46:41,955 - INFO - Creating /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV/tagmanifest-md5.txt
2018-10-20 06:46:42,013 - INFO - Creating /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV/tagmanifest-sha1.txt
(venv) [digilab@dl1-render-1 ZBOROV]$ echo $?
0
(venv) [digilab@dl1-render-1 ZBOROV]$ ls
bag-info.txt  bagit.txt  data  manifest-md5.txt  manifest-sha1.txt  tagmanifest-md5.txt  tagmanifest-sha1.txt

(venv) [digilab@dl1-render-1 ZBOROV]$
```

Automation to the rescue!

- Define inputs

- Implement steps by reusing existing tools and scripts

- Set validations and conditions for eventual branching

- Assign notifications

- Read the notifications :-)

- Audit the logs (GUI!)

There are generic tools available

- Jenkins
- Ansible
- Rundeck
- Buildbot
- ..
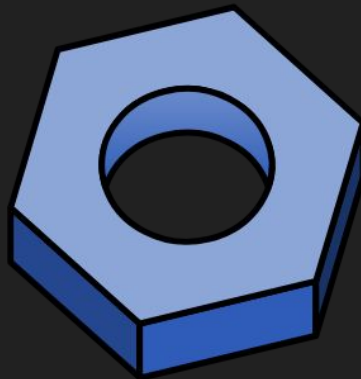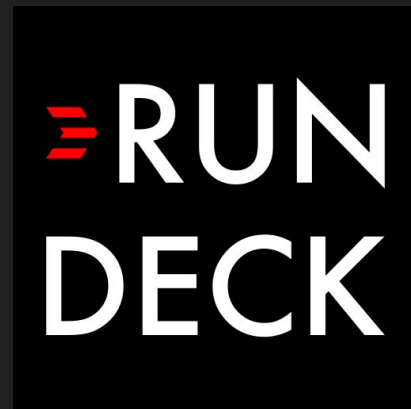
Jenkins

- Build automation server
- Web GUI, written in Java
- As it is generic, we can use it for whatever purpose
- Multi-tenant
- Plugin ecosystem

Jenkins jobs overview

Jenkins job parameters

Jenkins job parameters config

Jekins build log

Jenkins job tasks

Jenkins past build detail

Jenkins job chaining

Jenkins slaves

Jenkins slave tag matching

Jenkins as a cron

Jenkins job config in Git

Conclusions

- Standardized workflow needs deterministic steps
- As little human intervention as possible
- There are generic automation servers available
- Log retention!
- Jenkins is great

Thanks! Questions?

jonas.svatos@nfa.cz
github.com/NFAcz