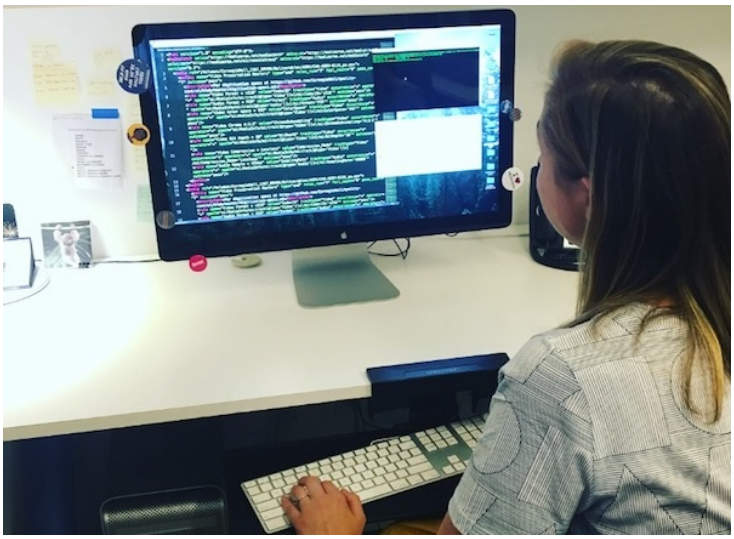
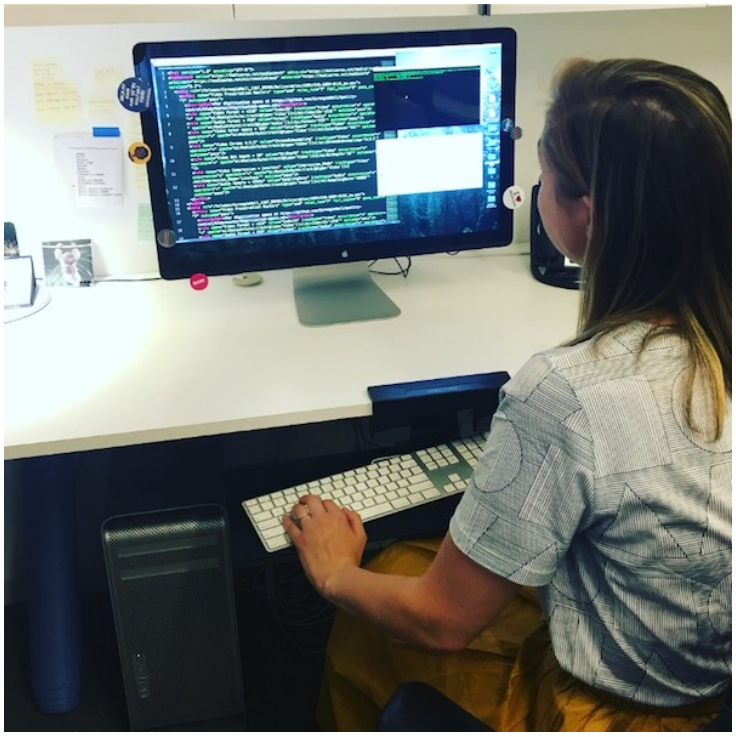
Interview with Kathryn Gronsbell





Kathryn with MediaConch

Hey Kathryn! Introduce yourself please.

Hi Ashley! I'm the Digital Collections Manager at Carnegie Hall. I develop and support sustainable practices around the digital asset lifecycle to ensure the availability and integrity of material related to the Hall and its history, collections, programs, and operations. I can be found talking about the struggling mass transit infrastructure in NYC, metadata quality assessment, and my Great Pyrenees rescue pup on Twitter.

What does your media ingest process look like? Does your media ingest process include any tests (manual or automated) on the incoming content? If so, what are the goals of those tests?

In 2012, the Carnegie Hall (CH) Archives started a multi-year initiative to digitize the majority of our physical holdings for preservation and access. We outsource our paper, video, audio, and film reformatting to different vendors and use a digital asset management system (DAMS) to organize, catalog, and present the material. Our quality control (QC) procedures for incoming digitized material are available on Carnegie Hall's Github. The process enables control and documented oversight from the point of hard drive / FTP delivery from a digitization vendor to ingest into our DAMS.

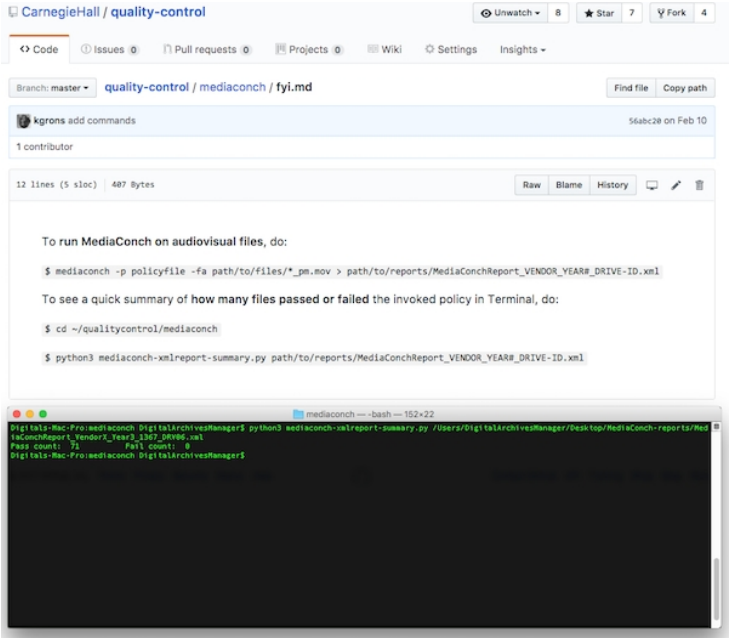
We aim to reduce risk while expediting the review and verification process with the QC procedures. The QC procedures increase our own accountability (How long does it take us to process 1 batch? What step is most time-intensive? Where can we expedite work by using different tools or workflows?) and allow us to better vet the continued work of our vendors (Are batches from the same vendor failing the same steps over time?). Another priority of the QC workflow is the ability to actually do it - the work is split between me and our Asset Cataloger, Lisa Barrier.

Where do you use MediaConch? Do you use MediaConch primarily for file validation, for local policy checking, for in-house quality control, for quality testing for vendor files?

Our primary use case for MediaConch is local policy checking against the digitization specs. We outsource our digitization, so quality testing the vendor output is a built-in function of our policy checking. We chose to balance manual review with automated testing: we perform manual visual/aural QC on 25% of a batch (or more, if the batch is small) and run MediaConch against the entire batch. I wrote a script which summarizes the MediaConch output to help expedite the review process for this step. We save MediaConch reports to an internal network drive for future use - we hope to build a digital repository in which we can submit submission information packages (SIPs) which contain information like the XML metadata from vendors and MediaConch reports.

At what point in the archival process do you use MediaConch?

MediaConch is part of Carnegie Hall's pre-ingest procedures.



Carnegie Hall Github page for MediaConch commands with terminal window

Do you use MediaConch for MKV/FFV1/LPCM video files, for other video files, for non-video files, or something else?

Our specs vary by source format, so we pass a variety of things through MediaConch (audio and video). We will be reviewing and likely revising our digitization specs in the next year, and MediaConch's ability to support Matroska, FFV1, etc. may play a role in our decision-making process.

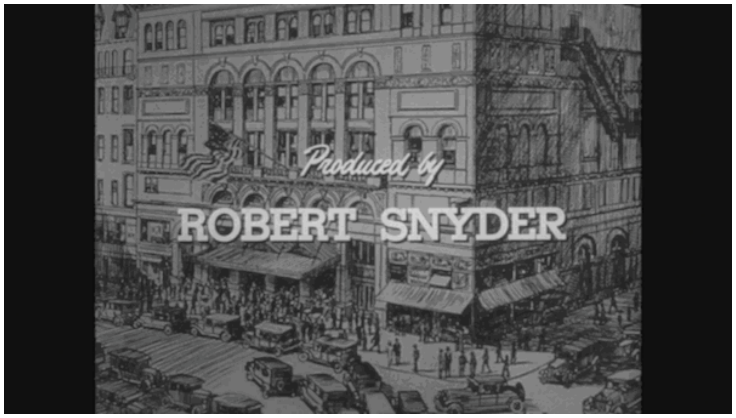
Why do you think file validation is important [or whatever you are doing?

There is an argument for verifying requested policies are being enforced on material digitized-as-a-service, and the ability to do so in-house, with a low learning curve. For my work in the CH Archives, I focus on how each step fits into the larger picture of what the entire workflow aims to achieve - accountability, dependability, and reproducibility.

Because of staff changeover and other factors, not all of our QC is completed in what I consider a ?normal' time frame. There is material digitized in 2013 that hasn't made it past half of the QC workflow (pause for screams of horror). In updating and improving our QC strategy, we acknowledge that the reality of our procedures mean batches may be processed asynchronously or in a wildly delayed timeframe, and manage those unfortunate symptoms of the prioritization juggling act that is preservation/archiving moving image material.

Anything else you'd like to add?

MediaConch is an incredible resource for Carnegie Hall. There was a few-month period where all QC on audiovisual files screeched to a halt. We were using MDQC, a policy checker built on top of ExifTool and MediaInfo, for audiovisual material. We ran into an issue which prevented us from analyzing files over ~200GB, despite a few months of troubleshooting. Many of Carnegie Hall's archival study recordings are full length concerts and performances, so we have some big uncompressed files to process. We still use MDQC for our still image material (concert programs, flyers, posts, photographs) but have transitioned to using MediaConch for any audiovisual material in the QC process. Without this tool, we would have a bigger QC backlog and would have needed to invest more money and time in determining how to facilitate policy checking audiovisual files. Thank you MediaConch creators, maintainers, and contributors!



?Concerts at Carnegie Hall - Michael Rabin, 1955? Michael Rabin, GIF Courtesy of the Carnegie Hall Archives

Thanks so much, Kathryn! For even more fun from the Carnegie Hall team, check out their Linked Data project repository.]