

On Evolving the Operations of the Vanderbilt Television News Archive—Cloud Computing for Audiovisual Preservation

Clifford B. Anderson

Abstract

This presentation describes the ongoing process of moving the operations of the Vanderbilt Television News Archive (VTNA) to the “cloud.” After briefly sketching the history and services of the VTNA, the author discusses how its staff have ported essential operations related to capturing, storing, transforming, and describing the evening news to cloud-based systems. This presentation was originally delivered on September 11, 2019 as part of a NISO webinar titled “The Preservation and Archiving of Digital Media.”

The focus of my talk today is the ongoing migration of the operations of the Vanderbilt Television News Archive to the “cloud.” I begin with a brief introduction to the history and services of the Archive and then proceed to discuss how and why we are transitioning our systems from our university computers and network to the third-party networks and services.

History

The Vanderbilt Television News Archive officially began on August 5, 1968. The origins of the Archive trace back to a partnership between Paul C. Simpson, an executive at the Metropolitan Life Insurance Company in Nashville, and Frank Grisham, Director of the Joint University Libraries at Vanderbilt University. As Simpson related the story in *Network Television News*, he turned on the evening news one evening in 1968 to discover Timothy Leary counseling the youth of America to “Tune in, turn on, and drop out.” Wanting to check whether he had misheard the interview, he requested a copy of the tape from the news network but he discovered to his dismay that no recording was available.¹ Simpson partnered with Grisham to preserve the evening news; Grisham employed students at circulation desk of the library to make the earliest recordings. Since its humble beginnings, the Vanderbilt Television News Archive has preserved half of a century of audiovisual data, documenting major cultural events including the

¹Paul Simpson, *Network Television News: Conviction, Controversy, and a Point of View* (Franklin, TN: Legacy Communications, 1995), 7.

moon landings, the final years of the Vietnam War, the Iran Hostage Crisis, the Space Shuttle Challenger disaster, the attacks of September 11, and beyond. Today, the VTNA operates with a staff of four full time professionals aided by librarians and staff members elsewhere in the Library who manage everything from its technical infrastructure to its business operations.

Services

The Vanderbilt Television News Archive provides three distinct kinds of service. First, the full collection of the Archive is available to any member of the public who visits our office on the campus of Vanderbilt University in Nashville, Tennessee. Over the years, researchers from disciplines ranging from history to media studies to political science and beyond have spent periods of time at the Archive, studying developments in the form and content of television news. Second, researchers can loan stories or shows from the entire collection. What we term a “compilation” consists of news clips, that is, one or more stories from news broadcasts. A “duplication,” by contrast, contains the entire broadcast, inclusive of commercials. The Archive charges a fee for these loans to recover its costs. Nowadays, the Archive generally loans media on USB drives, which researchers may keep for thirty days and then must return. A third service is available to nonprofit educational institutions in the United States. For sponsors of the Vanderbilt Television News Archive, the staff provides streaming access to its collection of news programming from CNN and NBC as well as discounts on loaned media.

Technology

The technology of the Archive developed over fifty years and, like an ancient city, different strata combine to support its surface architecture. While our collection is now almost entirely digital, the process of recording, capturing, describing, and providing access to its contents relies on a heterogenous set of technologies from distinct eras.

The workflow, at its simplest, is straightforward. The staff of the Archive records the evening news broadcasts of the major television networks and select cable news channels using an enterprise digital video recorder (DVR). They then apply a timeline to the recordings and create derivative versions, sending archival copies into longterm storage while generating lower resolution access copies for streaming. Next, staff members create metadata for the recordings, dividing the shows into segments and providing one or two sentence abstracts for each clip. The final step in the process is to link videos with metadata with shows and segments in a database, making the information about the shows and segments available to the public through our website.

This description is, of course, idealized. Exceptions exist to nearly every rule at the Archive. For instance, the staff members record breaking news as well as scheduled news. A DVR cannot foresee breaking events, meaning that we

need a strategy to capture unanticipated news broadcasts. We have burned a lot of tape (or hard drive space) saving six hour blocks of news programming to capture these events. As we develop stronger partnerships with other programs like the Internet Archive's Television News Archive, we can rest a little easier, knowing that if we miss an event, we should be able to retrieve a copy of the broadcast after the fact from a partner.

Capture Systems

Options for recording the evening news are myriad and seemingly expanding weekly. A primary distinction to make is between broadcast, cable, satellite, and over-the-top delivery systems. We are all familiar with broadcast systems, though the switch from analog transmission to digital signal in the United States and around the world introduced new advantages and complexities. Among these complexities is the difference in digital standard between the United States, which uses an Advanced Television Systems Committee (ATSC) standard, and Europe as well as much of the rest of the world, which uses a competing Digital Video Broadcasting Terrestrial standard. In some countries, cable has overtaken broadcast television and networks no longer broadcast television at all. In rural areas, direct-to-home satellite television still predominates. Over-the-top delivery, basically television programming delivered through the internet, is growing steadily.

The VTNA uses SnapStream to record the evening news broadcasts. SnapStream is an enterprise DVR capable of recording up to one hundred channels at once. The VTNA records from cable rather than over the air. As an aside, we capture the Nashville edition of the news, which may differ from broadcasts in other regions since local preemptions and commercial content may vary.

While the VTNA relies on enterprise technology to record the news, cheaper options exist. For instance, you can use a low-cost microcomputer like Raspberry Pi to capture television news programming. If you are unfamiliar with the Raspberry Pi, the Pi is essentially a complete, if low-powered, computer that comes without the typical peripherals. While designed for hobbyists, Raspberry Pis have found a wide range of business uses. The Raspberry Pi Foundation manufactures a series of so-called HATs (Hardware Attached on Top) that extend the functionality of the underlying Raspberry Pi, including a Raspberry Pi TV HAT to receive and stream digital television broadcasts.² The cost of this HAT is \$21.50, but it uses the DVB-T2 standard rather than the ATSC 3.0 standard. Since these devices do not have much computing power, the strategy would be to use them to capture broadcasts and to stream the data they collect to cloud-based servers for processing.

²See <https://www.raspberrypi.org/blog/raspberry-pi-tv-hat/>

Storage Systems

Another way that we rely on the cloud is to store our video files. After fifty years of recording, the collection of the VTNA takes up close to half a petabyte of storage. When you get to that level of data, standard options for storing the collection no longer function adequately.

The VTNA keeps its collection safe by storing copies in different formats and locations. We have full copies of the collection on videocassette and DVD, located offsite in a special collections facility, as well as on our supercomputing center, the Advanced Computing Center for Research and Education. A partnership between the Library of Congress and the VTNA also puts copies of every show we record on deposit at the National Audio-Visual Conservation Center (Packard Campus) of the Library of Congress in Culpeper, Virginia.

The best known cloud storage system is Amazon Simple Storage Service (S3).³ S3 operates like a file system in the cloud, allowing us to create “buckets” for video content. The cost of storing data in S3 depends on how often you need to access the data. Different options exist for standard, variable, and infrequent access. A related service called Amazon Glacier⁴ provides the ability to back up data to virtual tapes. The cost of keeping data on Glacier also reflects the likelihood of needing to get it back. If you need the data back, S3 Glacier allows you to retrieve it within the space of a couple of hours; S3 Glacier Deep Archive also allows you to reconnect with your data, but restoring it takes half a day.

AWS S3 serves as a foundational component in our emerging cloud-based architecture. We keep track of transformations in our data by moving videos, transcriptions, and other data from bucket to bucket in S3 according to a defined process, then depositing archival versions in S3 Deep Glacier for longterm storage.

Data Pipelines

At the core of the Vanderbilt Television News Archive is a workflow for recording, transcoding, describing, and making news programming accessible to the public. At some point, we would like to automate the entire end-to-end process. We are not there yet! But we have made significant progress toward that goal. At present, the process of moving videos and related metadata from SnapStream to AWS takes place according to an automated schedule. As soon as the video files land in an S3 bucket on AWS, they trigger so-called lambda functions to rename, transcode, and apply time codes.

AWS Lambda⁵ belongs to the category of “serverless computing”—that is, computing that does not require you to set up and maintain a web server that listens continually for API requests. By contrast, lambda functions execute in

³See <https://aws.amazon.com/s3/>

⁴See <https://aws.amazon.com/glacier/>

⁵See <https://aws.amazon.com/lambda/>

response to defined events within the Amazon Web Services ecosystem. These so-called “microservices” carry out a single action after such an event and then shut themselves down. The financial advantage of using lambda functions is that you pay when they fire, not while your system is idling. From a technical standpoint, using microservices allows us to reduce the complexity of our cloud computing environment by minimizing the scope of our code. If we need to update one of the lambdas, we do not have to worry about potential side effects elsewhere in our code base.

We use lambda functions to create a pipeline of transformations, ideally taking inputs from SnapStream and generating outputs for our streaming server through a series of discrete transformations. I say “ideally” because these lambdas do not always fire according to plan. More precisely, data does not always flow through the series of lambda functions as we expect. The staff of the VTNA maintain a spreadsheet to track successes and failures. Among the twelve error codes are “no recordings available in SnapStream,” “Exported from SnapStream, needs a manual lambda,” “AWS failed to convert,” “Corrected in AWS, download manually,” etc. The range of errors indicates the potential for the lambdas to go wrong. Right now, Steve Davis, the video archivist at the Television News Archive, logs these errors and handles exceptions manually.

Keeping track of such problems is a painstaking task and, to a degree, contradicts the goal of automating our data pipeline. On the one hand, the number of errors is decreasing as the staff learns to write more robust lambda functions to respond more flexibly to variances in the data. On the other, the orchestration of lambda functions is also possible through so-called AWS Step Functions.⁶ Step Functions provide a higher level framework for coordinating sequences of lambdas. In effect, you use them to create a finite state machine that keeps track of when lambdas succeed and what to do when they fail. We are not yet to the point of describing our environment entirely in the form of Step Functions, but we’ve begun to work in that direction and hope to supersede the manual logging within the near future.

Abstracts and Transcriptions

The website of the Vanderbilt Television News Archive provides distinct ways to explore and search its collection. If you are seeking for a news segment that you know or suspect aired on a specific date, you can pull up all the news programming from that evening and scan through the results for your segment. We also offer a keyword search with facets for date ranges, networks, reporters, and content type (i.e., news segments, commercials, special reports, etc.). To make this system work, the VTNA employs abstractors to divide news shows into segments, label them according to category, and to identify anchors and reporters. Crucially, abstractors write one or two sentence summaries of the content of the segments. When you search the website, the segments returned are

⁶See <https://aws.amazon.com/step-functions/>

those where the keywords match words and phrases in abstracts. Skip Pfeiffer, abstractor at the Archive, dedicates great care to encapsulating the essentials of every news story into his abstracts.

Reading through the abstracts provides succinct insight into daily news programming. But the system also has downsides for searchability. Patrons cannot search for specific words or phrases from segments as abstractors avoid quoting directly to avoid copyright infringement. Increasingly, evening news programming moves from topic to topic without defined intervals, making it difficult to break programs into segments. Anderson Cooper 360° on CNN, for instance, segues from issue to issue organically as its eponymous host moves the conversation along. Not to mention, divvying up a news show into segments is a laborious manual process.

For these reasons, some television news archives have given up segmenting news shows altogether in favor of intervals of fixed length. A search in these archives does not bring back a semantic segment, but indexes into specific intervals in the show by associating keywords with the video timecode.

As we move forward with modernizing the technical infrastructure at the Vanderbilt Television News Archive, we believe that providing semantic access to news segments represents an essential part of our DNA. But we also recognize the value of providing full-text search for the collection. We have begun to associate full text transcriptions of news segments with videos in our collection. Dana Currier, an abstractor at the Archive, has spent the past year exploring this new modality of describing the news.

We have considered three ways to generate these transcriptions. The most straightforward way would be to use the SubRip Subtitle (SRT) files we collect when recording shows. Those files are already synched with the timecode and enterprise services use them to provide full-text searchability. If you watch those subtitles, you know how frequently infelicities and errors creep into them. Another option is to generate transcriptions with Automatic Speech Recognition (ASR). Amazon offers a service with the novel name of Amazon Transcribe to provide textual transcriptions from audio and video at a reasonable cost. Using this service would be ideal for our purposes since we could then integrate transcription as another step in our finite state machine. Alas, our testing of the service did not get us to the level of accuracy we had hoped for. A final option was to produce our transcriptions with a computer-assisted transcription process. We looked at open source, commercial options, and the possibility of developing custom software. After conducting this review, we decided to select Trint⁷ for the task. Trint provides an online platform for AI-assisted transcribing and editing. The great thing about Trint is that it automates the first pass transcription and also lines up the video player and transcribed text according to their shared timecode. The transcriber uses a web interface to review the initial transcription, correcting and editing information as required while also

⁷See <https://trint.com/>

identifying segment divisions and enriching metadata.

With the generous support of The Peter & Carmen Lucia Buck Foundation, Jim Duran, Director of the Television News Archive, undertook a six month trial to transcribe a growing backlog of shows requiring abstracts in our collection. During the past months, five or so temporary workers have chipped away at that backlog, completing hundreds of transcriptions of news shows. Trint allows us to export the transcript in formats ranging from plain text to Word to XML, making it easy to integrate the results into our data pipeline.

Duran is also experimenting with using these transcripts to create semantic information about segments in the shows. For instance, he has conducted a successful trial with AWS Comprehend to extract entities from transcriptions.⁸ Comprehend is a natural language processing service that, among other services, identifies key phrases and topics in textual documents. By asking the transcribers to identify the start and stop times of the commercials in news broadcasts, Duran can send the transcribed text of these commercials to Comprehend to identify the topic. This system has proved effective apart from commercials where the product is never actually mentioned, apparently an increasingly prevalent advertising strategy.

Integrating these transcriptions into our discovery system requires us to change our user interface. As noted, we display the abstracts in their entirety because they are our interpretation of the news segments and, as such, our intellectual property. By contrast, the copyrights of the transcriptions belong to the respective networks. Our plan is to add them to our discovery in the form of a snippet view, showing keywords and phrases in the context of a sentence or two from their segments.

Machine Learning

As I wind up this presentation, I'd like to discuss the research questions we are investigating as we try to automate our archival processes. As you have heard me mention throughout the presentation, a distinguishing feature of the Vanderbilt Television News Archive is the division of news shows into semantic segments. As noted, this process takes place manually at present with abstractors and transcribers identifying the start and stop times of these segments. Ideally, we would use machine learning to identify the segments. This is an active area of computational research⁹ and the techniques can be complex, but the basic idea

⁸See <https://aws.amazon.com/comprehend/>

⁹See, e.g., A. G. Hauptmann and M. J. Witbrock, "Story Segmentation and Detection of Commercials in Broadcast News Video," in *Proceedings IEEE International Forum on Research and Technology Advances in Digital Libraries -ADL'98-*, 1998, 168–79, doi:10.1109/ADL.1998.670392, Andrew Rosenberg and Julia Hirschberg, "Story Segmentation of Broadcast News in English, Mandarin and Arabic," in *Proceedings of the Human Language Technology Conference of the NAACL, Companion Volume: Short Papers*, NAACL-Short '06 (Stroudsburg, PA: Association for Computational Linguistics, 2006), 125–28, and Z. Liu and Y. Wang, "TV News Story Segmentation Using Deep Neural Network," in *2018 IEEE International Conference on Multimedia Expo Workshops (ICMEW)*, 2018, 1–4,

is straightforward. We would build and train a model on a corpus of videos and/or textual transcriptions for which we have already identified and labelled the time codes of the segments. Our goal would then be to use that model to predict segment breaks in shows we have not yet described. As you might guess, there are cloud-based tools for supervised learning. On AWS, the relevant tool is Amazon SageMaker,¹⁰ which provides an essentially point-and-click framework for composing machine learning models.

Still, building a machine learning model is not trivial. While there are patterns in television news (such as the use of the nominative absolute when moving between stories, i.e. “Turning now to the weather...”) the shows exhibit variety both between themselves as well as diachronically. A question we also need to explore is whether we would need to train these models on the videos, transcriptions, or both. Given the complexities, we are exploring partnerships with faculty and graduate students in computer science to create and deploy machine learning models of this kind.

There remain “off the shelf” machine learning tools we hope to try out in the meantime. AWS Rekognition¹¹ provides a set of “out of the box” services to analyze videos, including celebrity recognition, scene detection, and the ability to pull out text from images. We hope that by using these tools in advance of our transcription efforts we can speed up our metadata work.

Challenges and Prospects

Let me conclude by listing challenges we are facing. First, we are seeking to keep pace with developments at both AWS and also other “clouds.” The major cloud environments—Amazon Web Services, Microsoft Azure, Google Cloud, and IBM Watson—are competing with one another to capture this fast-growing market. We try to pay attention to developments in these clouds, monitoring solutions and costs. For a long time, we relied on cloud storage at Oracle because the cost per file was lower than at Amazon. Taking a so-called “cloud agnostic” or “multi-cloud” approach inevitably requires working at a higher level of abstraction and expertise in third party frameworks like the Serverless Framework.¹² In the end, the limitations of human resources at Vanderbilt Television News Archive lead us to look to Amazon Web Services before considering competing frameworks. If we cannot find a solution at Amazon, we frequently find that other software companies integrate their products with Amazon Web Services.

The shift to the cloud has been positive for the Vanderbilt Television News Archive. We have made gains in reliability, insight, and operational agility. We see a wide horizon of opportunities to add value to our data by drawing on existing and emerging cloud-based tools. But this shift constantly stretches us, pushing us to learn new IT frameworks, programming languages, and protocols.

doi:10.1109/ICMEW.2018.8551568

¹⁰See <https://aws.amazon.com/sagemaker/>

¹¹See <https://aws.amazon.com/rekognition/>

¹²See <https://serverless.com/>

We are keenly aware that we cannot make this transition alone, but must work in partnership with IT colleagues in the library, the university, and at commercial firms. In the long term, the biggest change may not be porting our operations to the cloud, but the accompanying changes to our work culture as we embrace a DevOps¹³ mindset: working in partnership with IT specialists to improve our systems collaboratively step by step, bit by bit as we seek to meet the evolving needs of our patrons.

Works Cited

Ebert, Christof, Gorka Gallardo, Josune Hernantes, and Nicolas Serrano. “DevOps.” *IEEE Software* 33, no. 3 (May 2016): 94–100. doi:10.1109/MS.2016.68.

Hauptmann, A. G., and M. J. Witbrock. “Story Segmentation and Detection of Commercials in Broadcast News Video.” In *Proceedings IEEE International Forum on Research and Technology Advances in Digital Libraries -ADL’98-*, 168–79, 1998. doi:10.1109/ADL.1998.670392.

Liu, Z., and Y. Wang. “TV News Story Segmentation Using Deep Neural Network.” In *2018 IEEE International Conference on Multimedia Expo Workshops (ICMEW)*, 1–4, 2018. doi:10.1109/ICMEW.2018.8551568.

Rosenberg, Andrew, and Julia Hirschberg. “Story Segmentation of Broadcast News in English, Mandarin and Arabic.” In *Proceedings of the Human Language Technology Conference of the NAACL, Companion Volume: Short Papers*, 125–28. NAACL-Short ’06. Stroudsburg, PA: Association for Computational Linguistics, 2006.

Simpson, Paul. *Network Television News: Conviction, Controversy, and a Point of View*. Franklin, TN: Legacy Communications, 1995.

¹³Christof Ebert et al., “DevOps,” *IEEE Software* 33, no. 3 (May 2016): 94–100, doi:10.1109/MS.2016.68