



NEH Grant #
PW-259067-18

Submitted on
August 31, 2021

WHITE PAPER

Unlocking the Airwaves:

*Revitalizing an Early Public and
Educational Radio Collection*

Principal Investigators

Stephanie Sapienza

Digital Humanities Archivist
University of Maryland (MITH)

Eric Hoyt

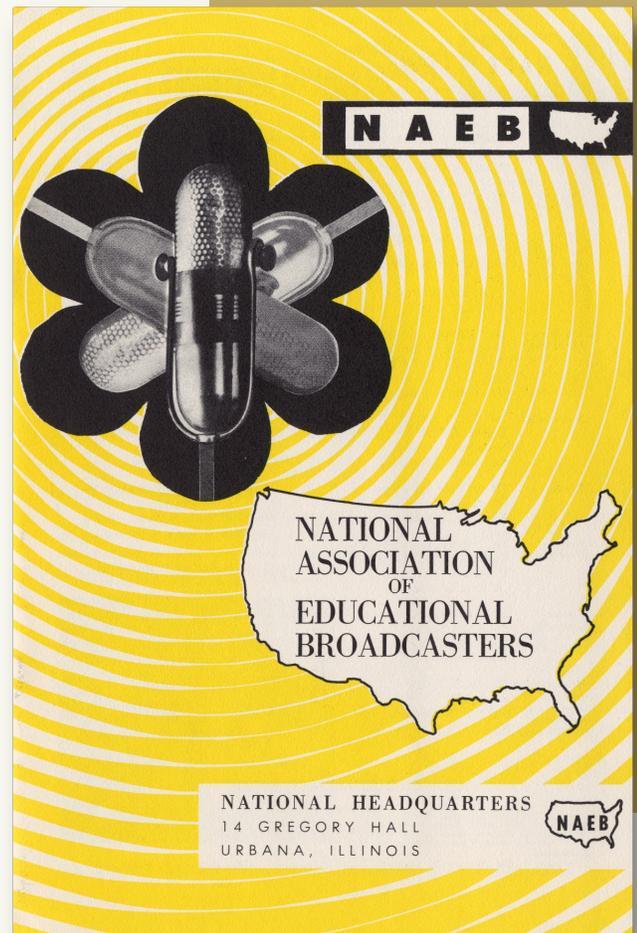
Professor of Media & Cultural Studies
University of Wisconsin-Madison

Introduction

Anytime a listener downloads the TED Radio Hour podcast to her iPhone, or turns on the car radio and hears Terry Gross interviewing an author on Fresh Air, that listener is participating in a 100-year-old tradition of utilizing audio broadcasting to educate and inform the American public. The pivotal, yet largely forgotten, institution responsible for promoting educational and noncommercial broadcasting for most of the 20th century was the National Association of Educational Broadcasters (NAEB). From the mid-1920s until the early-1980s, the NAEB developed and distributed educational audio programs and accompanying print materials to schools and communities across the United States. What's more, the NAEB lobbied extensively to unlock the airwaves—to access precious frequency space—in order to bring the voices of poet Robert Frost, architect Frank Lloyd Wright, anthropologist Margaret Mead, conservationist “Ranger Mac,” and many other individuals into American homes and classrooms. Before NPR and our current “public radio,” there was the NAEB and “educational radio.”

The NAEB's history is the dramatic story of idealists who believed in the utopian possibilities of technology for education and social uplift and who faced considerable challenges along the way, including economic depression, world war, and the scarcity of the electromagnetic spectrum. It's a story that has much to tell us about 20th century American society; the topics covered in NAEB recordings are as diverse as the national census, atomic energy, American labor, religion, US history, agricultural engineering, mathematics, and foreign relations.

Unlocking the Airwaves: Revitalizing an Early Public Radio Collection sought to realize the potential of the collections of the National Association of Educational Broadcasters (NAEB) for exploration and study by educators, scholars, journalists, documentarians, genealogists, and the broader public. By coordinating the expertise of archivists, humanities researchers, and digital humanists, we wanted to create a new online resource for humanities research that would deliver enhanced access to important, mostly hidden, archival audiovisual materials.



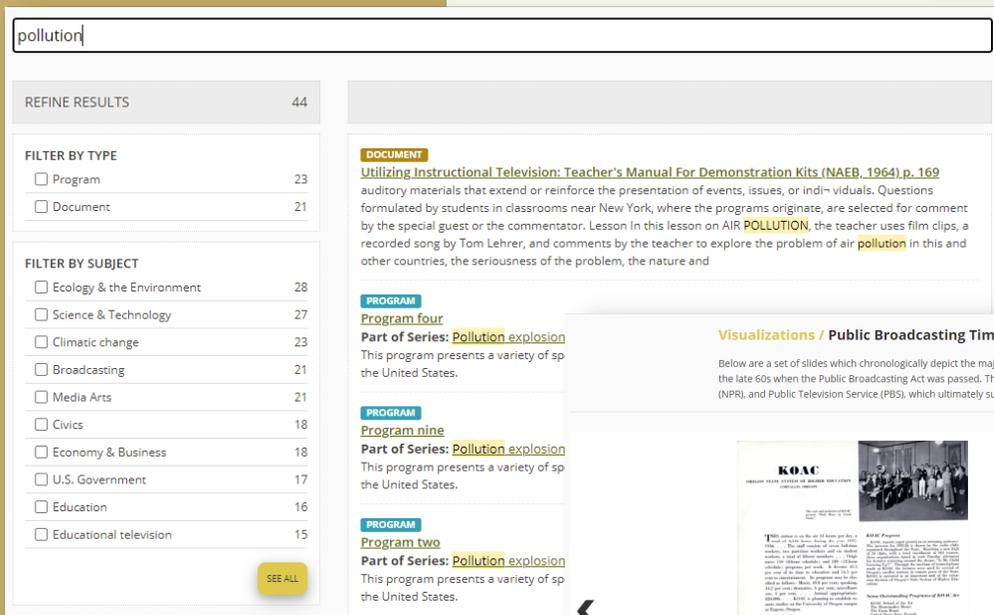
Introduction (continued)

VISUALIZATIONS

Three visualizations providing unique modes of exploring the collections, including an interactive map of content creators; a timeline of public broadcasting history; and an interactive tree diagram linking educational content standards to specific resources in the collections.

SEARCH

A search platform for the linked collections, enabling users to simultaneously search both the documents and sounds of the NAEB.



(Left) The search interface on the Airwaves website; (Below) An interactive timeline visualization on the Airwaves website.



FRAMEWORK

Implementation of an innovative framework to do this work in a way that best utilizes the potential of each standard while also enhancing the interoperable functionality of the others.

Project Origins (continued)

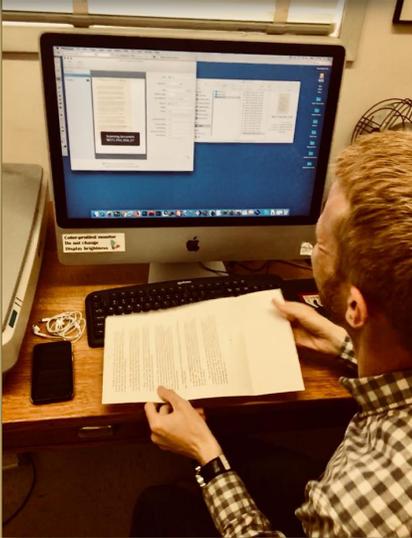
Eventually, the team included Co-PI Eric Hoyt at the University of Wisconsin-Madison, due to his knowledge of the subject matter and his groundbreaking work designing an interface for exploring media-related documents at the Media History Digital Library (mediahistoryproject.org), and Josh Shepperd at Catholic University, whose research interests were directly aligned with the project. Professor Shepperd was at the time the Director of the Library of Congress's Radio Preservation Task Force (RPTF). The RPTF had the potential to be an avenue for collaboration on outreach activities, since it already had plans to engage educational partners in developing curricular materials for the study of radio.

Research Questions

At the heart of this project lies a question that both Principal Investigators have long identified as being central to many ongoing conversations surrounding the role of the audiovisual in digital humanities: ***how can we utilize digital humanities methods to encourage the use of audiovisual collections as primary historical records, both in and outside of the academy?***

The project team published a lengthy piece which unpacked the many complexities of this research question - why these collections become marginalized in the first place, how their descriptive strategies present unique challenges in a number of directions, and what other challenges face digital humanists who want to use a/v collections. Rather than repackaging or summarizing all of those complexities here, we encourage those interested to access that article in the 2021 Volume 15 edition of *Digital Humanities Quarterly*. Below, we'll instead concentrate in the final sections of this white paper on the final outcomes of the project, as well as the various ways its impact and findings can be seen and experienced.

(Below Middle) The Airwaves project team at the Wisconsin Historical Society - left to right: Stephanie Sapienza, Paul Hedges, Matt St. John, Eric Hoyt, J.J. Bersch; (Below Bottom) Matt St. John digitizing materials from the NAEB paper collections.



Project Impact & Outcomes

EXPANDED ARCHIVAL WORKFLOWS AND TRAINING FOR AIRTABLE

Towards the beginning of the *Airwaves* project, Sapienza became acquainted with Airtable through the work of Miriam Posner on a project on early race film at UCLA, and immediately recognized that the tool afforded possibilities and expanded functionality for data-driven projects that many cultural stewards struggle to overcome. The cloud-based relational database could handle one to many relationships; it could store controlled vocabularies and taxonomies in separate linked tables; the Pro version contains apps for batch update and de-duplication actions; and custom views can be created for subsets of a collection pertaining to a particular research need or request. Airtable is managed through any browser and is easy to use – it looks like Excel, but functions better than FileMaker Pro and OpenRefine combined.

Over the course of the project's completion, Airtable's merits have never stopped growing, at times exponentially. Now it is the backbone of many aspects of the project – not only did the project team use it to manage our sustainability planning, digitization tracking, and various project management tasks, but it has actually become the backend for a very complex data model behind the *Airwaves* site. This acquired knowledge has led to countless other outcomes.

A workshop, *Airtable for A/V Archivists*, led by Sapienza and Ed Summers at the Association of Audiovisual Archivists in 2019 was so popular that it was repeated during the pandemic as a webinar for AMIA's Continuing Education Advocacy Task Force for 100+ participants.

(Below Middle) Airtable base tracking digitization progress; (Below Bottom) Presentation title slide from the Airtable for AV Archivists workshop in 2019..

Folder	ID	Folder Name	Assigned to	Priority	Progress	Internet Archive URL	
Completed	107	naab-0112-001	naab-0112-001	112	Newsletters, Newsletter (general membership), 1965-1967	Completed	https://archive.org/details/naab-0112-001
	108	naab-0112-002	naab-0112-002	112	Newsletters, President's Memo, 1954-1966	Completed	https://archive.org/details/naab-0112-002
	109	naab-0112-003	naab-0112-003	112	Newsletters, President's Memo, 1954-1966	Completed	https://archive.org/details/naab-0112-003
	110	naab-0112-004	naab-0112-004	112	Newsletters, Television Engineering, 1955-1957	Completed	https://archive.org/details/naab-0112-004
	111	naab-0112-005	naab-0112-005	112	Newsletters, Washington Report, 1960-1962	Completed	https://archive.org/details/naab-0112-005
	112	naab-0112-006	naab-0112-006	112	Newsletters, Washington Report, 1960-1965	Completed	https://archive.org/details/naab-0112-006
In Progress (Metadata)	113	naab-0101-001	naab-0101-001	101	Publications, 1920-1939	In Progress	https://archive.org/details/naab-0101-001
	114	naab-0101-002	naab-0101-002	101	Publications, 1920-1939	In Progress	https://archive.org/details/naab-0101-002



Project Impact & Outcomes (continued)

Sapienza, along with Purdom Lindblad, taught Airtable three separate times to graduate students in MITH’s Anatomy of Digital Humanities course. Airtable has now heavily informed an even more complex data model for another concurrent NEH-funded MITH project, the Lakeland Digital Archive (Grant #s PY-263763-19 and HAA-269068-20). Sapienza has done at least five consultations on how to implement an Airtable base for projects both tiny and massive, including most recently for the folks working on the The Baltimore Afro-American newspaper archives, who are migrating a massive collection out of Archon and want to digitize thousands of pages and eventually host them online. MITH is now using Airtable as the backend of its own website. All of these innovations and extensions were made possible by the initial use of Airtable for *Unlocking the Airwaves*.

ID	location	file cabinet	shelf	container	folder title	physical condition	condition	type
1	Hall:01:01	01	01	001	Adams, Little Willie	Fair Condition: 2 folders; 8 items: photographs, clippings; Location: Hall...	Fair	Photographs Clippings
2	Hall:01:01	01	01	002	Accident Victims	Type: clippings, photographs; ; Box Number:: ; Folder Number:: ; Locati...		Photographs Clippings
3	Hall:01:01	01	01	003	Academy Awards	Type:: ; Box Number:: ; Folder Number:: ; Location:: ; ; Lists of Award ...		
4	Hall:01:01	01	01	004	Abbott, Robert Sengstacke	Type: clippings, court documents, notes, photographs, press releases, p...		Photographs Clippings
5	Hall:01:01	01	01	005	Afro Building Pix [Pictures]	Fair Condition: 11 items: Photographs; Location: Hall:01:01	Fair	Photographs
6	Hall:01:01	01	01	006	Afro, Hoe Straight-Line Press	Type: photographs; Box Number: Folder Number: Location: Hall:01:01		Photographs
7	Hall:01:01	01	01	007	African Princess, Hadiza Coujেকেe N...	Fair Condition: 3 items: Photographs; Location: Hall:01:01	Fair	Photographs
8	Hall:01:01	01	01	008	African Pictures, AID	Fair Condition: 10 items: Photographs (publicity), press releases; Locati...	Fair	Photographs
9	Hall:01:01	01	01	009	African Methodist Episcopal Church (...)	Fair Condition: 6 items: Clippings, programs, notes, correspondence; Lo...	Fair	Clippings
10	Hall:01:01	01	01	010	African Massacre	Fair Condition: 1 item: clipping; Location: Hall:01:01	Fair	
11	Hall:01:01	01	01	011	Ades, Bernard	Fair Condition: 2 items: Hall:01:01	Fair	
12	Hall:01:01	01	01	012	AFRO, 1944 Edition	Fair Condition: 4 items: clippings ; Hall:01:01	Fair	Clippings
13	Hall:01:01	01	01	013	AFRO, Clean Block Campaign (3 fold...	photographs, clippings; Hall:01:01		Photographs Clippings
14	Hall:01:01	01	01	014	AFRO Cooking School	Fair Condition: 2 items: photographs; Hall:01:01	Fair	Photographs
15	Hall:01:01	01	01	015	AFRO, Lucky Buck Contest	Good Condition: 1 item: photograph; Hall:01:01	Good	
16	Hall:01:01	01	01	016	AFRO, 75th Anniversary	Fair Condition: 2 items: clippings; Hall:01:01	Fair	Clippings
17	Hall:01:01	01	01	017	AFRO, 100th Anniversary	Good Condition: 1 item; program; Hall:01:01	Good	
18	Hall:01:01	01	01	018	AFRO, Financial Report, 1949	Good Condition: 1 item; report; Hall:01:01	Good	
19	Hall:01:01	01	01	019	Afro Employees, Circulation	Good Condition: 2 items: photographs; Hall:01:01	Good	Photographs

Airtable prototype demo created for the Baltimore Afro-American newspaper archives. The model for the prototype was the Documents table in the Airwaves base.

Project Impact & Outcomes (continued)

ADVANCING TECHNIQUES AND APPROACHES FOR HUMANITIES REFERENCE RESOURCES USING STATIC WEBSITES

One of the innovative aspects of the *Unlocking the Airwaves* website is that it offers a dynamic website experience yet runs using static website technology. The advantages of the static site approach for building digital projects devoted to cultural heritage collections are numerous. Building and supporting databases and search indexes devoted to individualized projects are cumbersome and expensive to maintain. So, too, is relying on proprietary and subscription-based content management systems. In contrast, by using a static site generator, it is possible to create a website that feels highly dynamic to the end user but that is served up as client-side files within a conventional directory structure (HTML, CSS, JavaScript, images).

The *Airwaves* website holds relevance for other digital projects in the field. It was built using Gatsby, a static website generator written in NodeJS. Even though the website is deployed as a set of static files to a web server, it relies on several other services during its build and runtime. The build is a process that happens once when the site is deployed, and the runtime is the process run by your browser when you access the website. These services are listed in detail in the site documentation in the [GitHub repository](#). Additionally, we produced a short video for use in training and outreach, which describes in simpler terms how this process works. [That video can be viewed here](#).

We are not aware of any other projects in the field that allow so much searchability and adopt this static site approach, rather than relying upon querying a database such as MySQL. The use of Airtable allows the site to utilize a very complex data model, but that complexity doesn't translate to unwieldy server-side dependencies.

Project Impact & Outcomes (continued)

This said, one major finding is that we may have ‘hit the limit’ of this complexity. Because we wanted the site to look and feel in many ways like a traditional user interface with facets, keyword search, etc (like websites, for example, which employ a Solr index or similar), we ended up with a very large index file which loads into the browser. So far this large index file has never crashed or caused any issues with search, but it is definitely larger than what would ideally be loading into a static site of this flavor. So, while there are certainly trade-offs to the static site approach, we believe that *Unlocking the Airwaves* offers an exciting proof-of-concept for a more sustainable approach to building collection-oriented websites for libraries, archives, and other heritage organizations.

LOWERING THE BARRIER TO ENTRY FOR LINKED DATA PROJECTS

With linked data, we have the potential to realize the semantic web’s dream of broad discoverability and interoperability. But for institutions to realize that potential requires a fundamental shift from “thinking about metadata as human-readable records to thinking about it as machine-readable entities.” ‘Triplestore’ databases and data modeling specifications such as the Resource Description Framework (RDF) are historically two of the defining characteristics of linked data projects. But there are several challenges that lead to barriers for implementation by many cultural stewards. Triplestore databases typically have to be custom-built with a high technical barrier to entry, and often do not have the kind of easy querying capabilities of a typical database, nor do they have user interfaces which are understandable by the average cultural steward. Data has to be imported and exported using RDF.

The quotation above is from the 2015 “IMLS Focus Summary Report: National Digital Platform” report. As that report summarized, more institutions need simple tools and models to follow. *Airwaves* is a prime example of a model project that makes linked data principles more approachable and understandable.

Project Impact & Outcomes (continued)

Airwaves uses the International Image Interoperability Framework (IIIF) to display the scanned NAEB paper documents via a Mirador viewer on the site, which are hosted at the Internet Archive. This is linked data. The final enhanced metadata is published as JSON-LD and sent back to UMD Libraries and the American Archive. That is also linked data. The landing pages on the website for entities (People and Organizations) represented in the collections were developed after a process of reconciling the names to Wikidata and the Social Networks and Archival Context (SNAC) databases, and the URLs for those entity pages use the Wikidata Q-item identifiers. That is also linked data. And the *Airwaves* team did not maintain one triplestore database on this project.



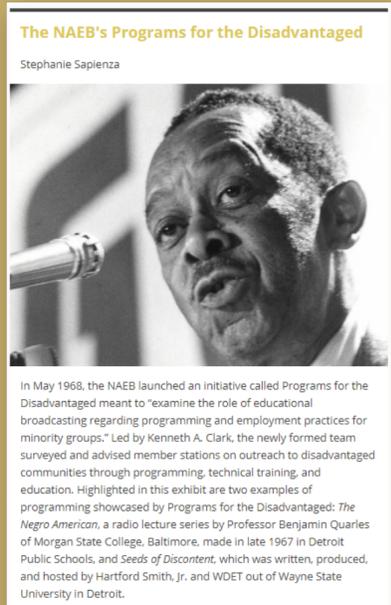
In January 2021, the NEH generously awarded MITH and UW-Madison funding for an *Airwaves* 'sequel' project (of sorts), specifically about how to develop a linked data framework using Wikidata and SNAC, in order to increase understanding and contextual associations between disparate radio collections. That project, *Broadcasting Audiovisual Data*, has already begun in earnest, and we've hired a new Graduate Research Assistant, Emily Frazier. In addition to the tutorial videos and workflows we have created for *Airwaves* (which are all linked to in the Appendices section), this next project will create additional unambiguous video tutorials on topics such as how to take a flat dataset, extract archival authority data, and use OpenRefine to reconcile and pull data from the web. To do this using only human labor may take several full-time catalogers one or more years. Therefore, our findings and the products we created out of them are helping to enrich resources on the web while saving archives hundreds of man-hours.

Project Impact & Outcomes (continued)

ACQUISITION OF THE HARTFORD SMITH, JR. COLLECTION

In January 2021 the *Airwaves* team was contacted by a woman named Chana Smith. Her father, Hartford Smith, Jr., produced the late-60s radio series *Seeds of Discontent*, which is part of the *Airwaves* collections and is featured on the website. Mrs. Smith connected the team with her father, who conducted a virtual interview which is now linked to from an [exhibit on the site about the 'Programs for the Disadvantaged' initiative](#). Professor Smith also informed the team that he had all the original audio tapes from the interviews he conducted for the series, as well as the missing first episode that the collections were missing. Over the course of the next six months, a deed of gift was drafted and signed between Professor Smith and the Wisconsin Center for Film & Theater Research (WCFT) at the University of Wisconsin-Madison, under the direction *Airwaves* Co-PI Eric Hoyt. The collection was inventoried and is now in the hands of the WCFT, soon to be digitized and transcribed.

Seeds of Discontent was unique in the fact that it featured Studs Terkel-style 'man on the street' interviews with average people, at a time when a lot of educational radio content was highly controlled and studio-produced. It also immediately followed the Detroit riots, and featured black citizens of Detroit speaking their minds comfortably and candidly. For these and other reasons, this new complementary collection is very culturally important. And it may not have ever surfaced if it weren't for the visibility of the *Airwaves* website.



(Above) *The Programs for the Disadvantaged* exhibit on the *Airwaves* site.



The Title for the first 1/2 hour installment is
 Subject: "The American Negro: The Historical Perspective"
 The Negro will create
 the perspective on riots, rebellion, and the
 displacement of Negro citizens through
 the eyes of older Negro citizens who have
 lived in Detroit since the turn of the
 century.
 It will stress the
 continuity of this experience & history
 in the young as they relate to the
 living history experienced through their
 elders.
 The recording enclosed demonstrates
 the documents already collected for
 this specific production.

(Above) Sapienza inventories the Hartford Smith, Jr. collection in Detroit; (Below) Items in the Hartford Smith, Jr. Collection

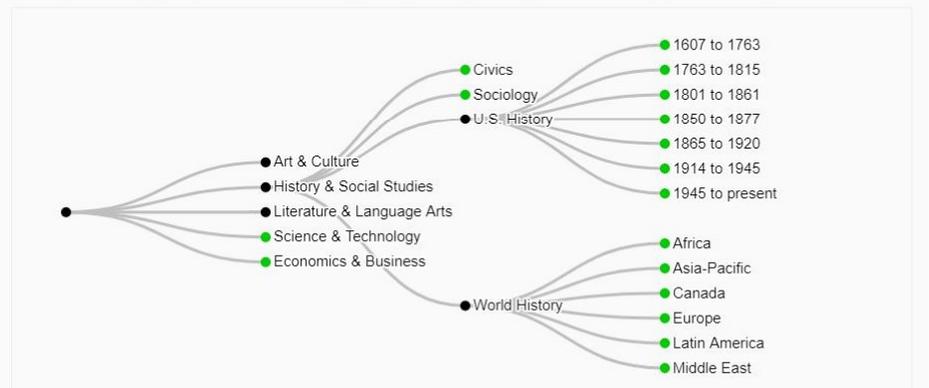
Project Impact & Outcomes (continued)

NEW INSIGHTS ABOUT USING AUDIOVISUAL MATERIALS IN CLASSROOM INSTRUCTION

In July 2021, in conjunction with the Radio Preservation Task Force Education Division, Sapienza worked with The History Project, a professional development organization at the University of California Irvine, to sponsor a digital storytelling workshop for primary and secondary school educators interested in integrating more audiovisual media into their classrooms. The workshop provided a general overview of the value of digital storytelling for meeting diverse educational goals—from presenting new material and alternative assignments to supporting students’ digital literacy. It also specifically aimed to measure the effectiveness and impact of the content standards visualization on the Airwaves site, which maps the radio programs and documents in the Airwaves collections of educational content areas using ontologies from EDSITEMent and the National Assessment of Educational Progress.

Visualizations / Educational Content Areas Tree

This interactive tree maps the radio programs and documents in Unlocking the Airwaves to educational content areas using ontologies from EDSITEMent and the National Assessment of Educational Progress. Click any black dot to expand or collapse the tree and see content areas related to that parent topic. Click any green dot to open a link in a new tab and explore the radio programs and other materials on Unlocking the Airwaves related to that content area.



Content Standards Visualization on the Airwaves website

Project Impact & Outcomes (continued)

The second half of the workshop consisted of several focus groups asking participants about their interest and experience with audiovisual media in the classroom. Through this process, we learned that educators in primary, secondary, and museum settings are interested in utilizing more audio. Many said that the major impediment to doing such work was a lack of awareness of their existence, combined with limited access to digitized or transcribed materials.

The full report on these findings can be found in [this blog post](#). But in short, we discovered that teachers need 1) short audio segments 2) thematically linked to each other and to visuals and other media 3) with discussion questions or an outline 4) especially in the areas of marginalized voices, incl. languages other than English, 5) linked to curricular standards. For student-produced work, many are looking for tutorials that students themselves can use, rather than teaching production as a major class component themselves.

(Right) Prezi slide deck for the workshop;
(Below) List of resources made available to educators. Links to both can be found in the Appendices section.



Name	Resource Category	Type	URL	Notes	Status
Archival Materials					
Storytelling Tools (Count: 16)					
Tutorials (Count: 18)					
53	Digital Storytelling Using Audio Sources A Radio Preservation Task Force (RPTF) Workshop	Tutorials	Training	https://prezi.com/view/Dm5CZbc02Ww434etkva/	This is the Prezi presentation from the workshop!
54	Digital Storytelling Sizzle Reel: Guide to Clips Used	Tutorials	Sample Student Work	https://airtable.com/shrXEO3Zca0H4cd0s	This is the guide to all clips used in the 'Sizzle Reel' for the workshop. The edited reel itself (also in the Prezi link above) is on Vimeo at https://vimeo.com/56858530 .
55	Skillshare	Tutorials	Training	https://www.skillshare.com/	Nerdwriter is offering a special promo of two free months of Skillshare Premium. visit this link!
56	How to Record Your Interviews in Zencast	Tutorials	Video	https://www.youtube.com/watch?v=kX98TeU4Hw	Guide to recording interview videos virtually using Zencast.
57	Guest Video: Record Video, Record Audio	Tutorials	Video	https://www.youtube.com/watch?v=Vfo3ayCDng	Guide to recording interview both audio and videos virtually using Zencast.
58	Searching newspaper articles from specific countries	Tutorials	Research & Organization	https://www.youtube.com/watch?v=5Eidkw5FMb0	Guide to finding and using newspapers in student projects.
59	10 Sites for Free Stock Photos	Tutorials	Research & Organization	https://www.fileon.com/blog/best-free-business-stock-photo-websites	A guide to finding free resources for stock photos (stills) to use as B-roll.
60	Documentary & Stills Mind Map of Ideas	Tutorials	Research & Organization	https://alenigeorgiouphotography.wordpress.com/2017/01/25/documentary-stills-mind-map-of-ideas/	Idea map of how to use stills to convey meaning in digital storytelling.

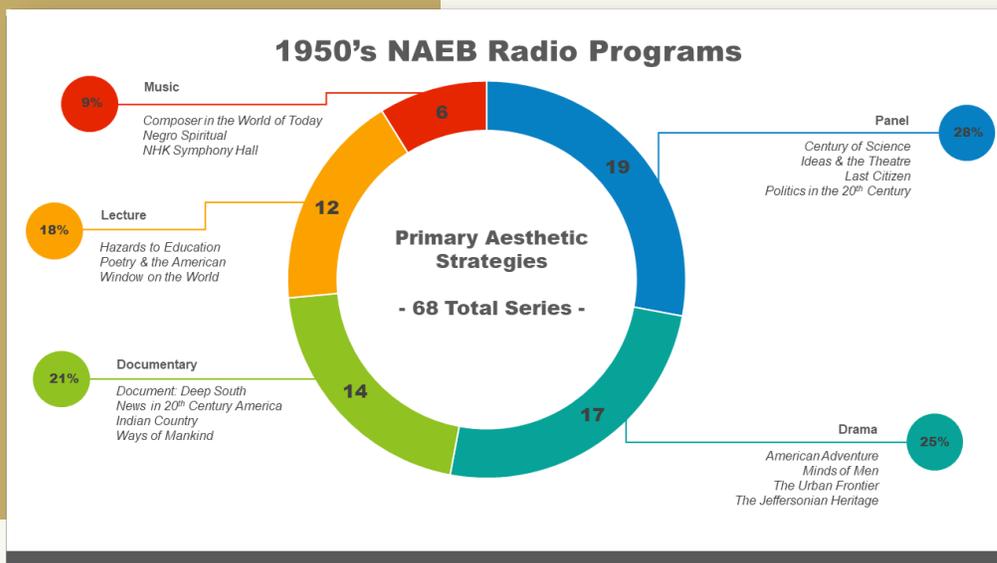
Project Impact & Outcomes (continued)

RESEARCH CONCLUSIONS ABOUT EARLY PUBLIC RADIO'S AESTHETIC STRATEGIES

PI's Hoyt and Sapienza, along with Production Assistant Matt St. John, completed a draft chapter for the forthcoming *Oxford Handbook of Radio Studies* book. The chapter is about the history of production norms and aesthetic strategies in public radio, tracing backward from our current moment of podcast ubiquity and public radio dominance, and weighed against popular notions about so-called 'NPR Voice' and its 'carefully casual' white hegemony. To bolster and contextualize these arguments, we walked back through and unpacked the nuances of the NAEB's entire program output, and attempted to paint a more complete picture of how radio's aesthetic strategies have always been widely varied, paralleling major changes in U.S. history, politics, and culture.

This kind of work was only made possible due to the presence of the *Airwaves* resource, which grants unprecedented access to the largest corpus of early educational radio ever made publicly available. One of many reasons this was a rewarding culminating experience is that Matt St. John has been involved in the *Airwaves* project since the beginning, and had personally (alongside J.J. Bersch) digitized and described

hundreds of documents at UW-Madison, linking them to radio programs. His knowledge of the collections in many ways exceeded that of anyone else on the project team, and our combined perspectives on these collections led us to new and interesting ways to frame the chapter. *The Oxford Handbook of Radio Studies* will be published later this year, and will be publicized by the project team on the website and elsewhere.



Graphic illustrating the various aesthetic strategies in NAEB programs

Appendices

WEBSITE DOCUMENTATION

[GitHub Documentation](#)

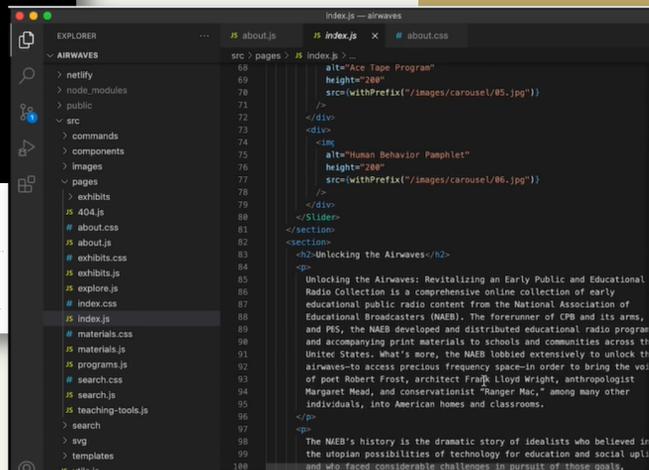
In addition to containing all the publicly-accessible site code, the Read Me file at the Airwaves GitHub repository contains extensive documentation about the build and runtime environment employed by the website.

[The Unlocking the Airwaves Website: A Brief Guide](#) (video)

This short video employs complementary visuals to describe in simpler terms how the *Airwaves* website works. The video was designed for a less technical audience, and can be utilized in classrooms or workshops.



Screenshots from the Airwaves website demonstration video described above.



DIGITIZATION WORKFLOW

[Unlocking the Airwaves: Digitization Workflow Guidelines](#)

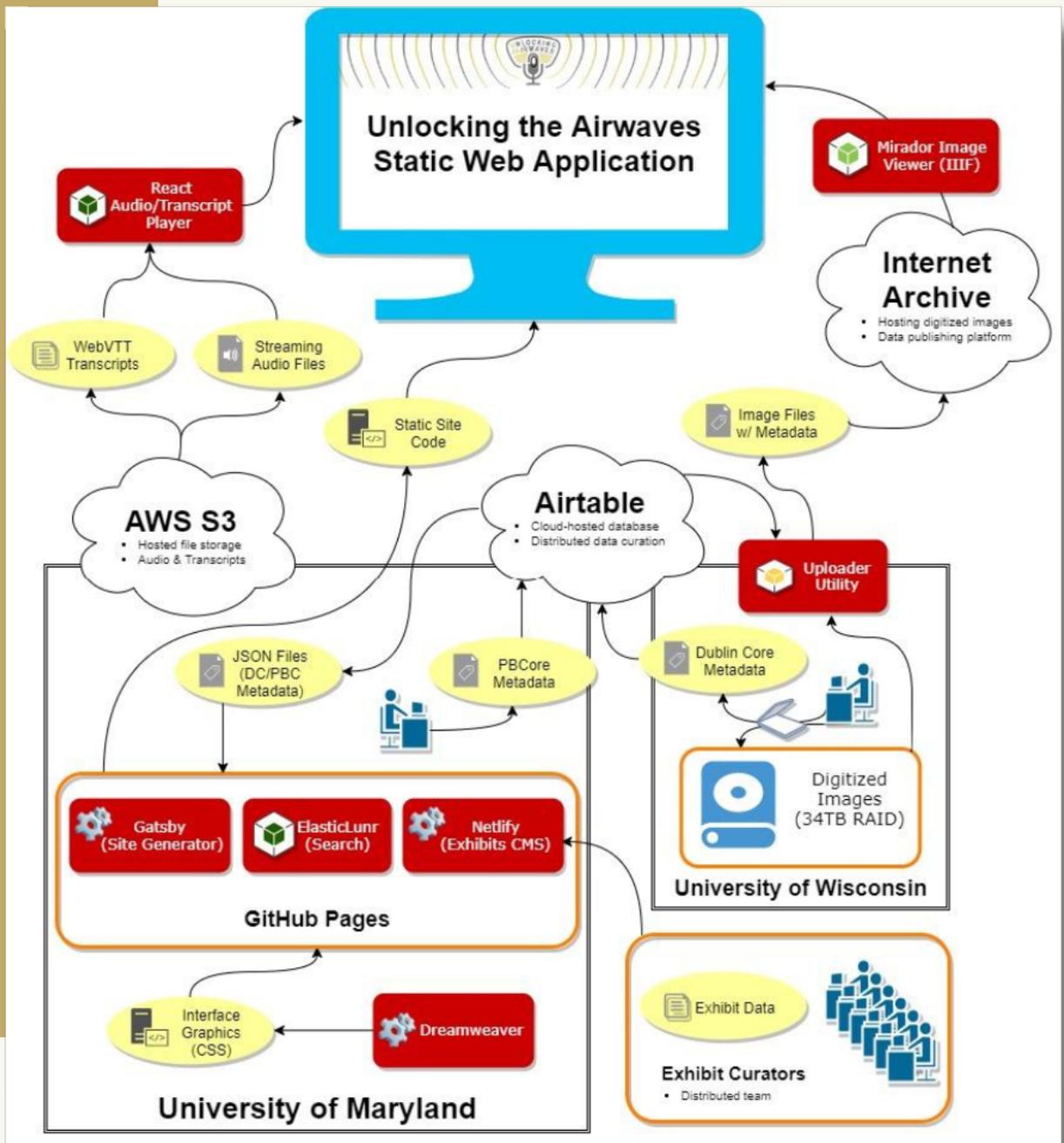
This document gives a detailed overview of our procedures for digitizing the paper collections of the NAEB at the University of Wisconsin-Madison, including technical specifications, color corrections, metadata capture, and uploading materials to the Internet Archive.

Appendices (continued)

DIAGRAMS, MODELS, AND GRAPHICS

[System Design and Process Flow Diagram](#)

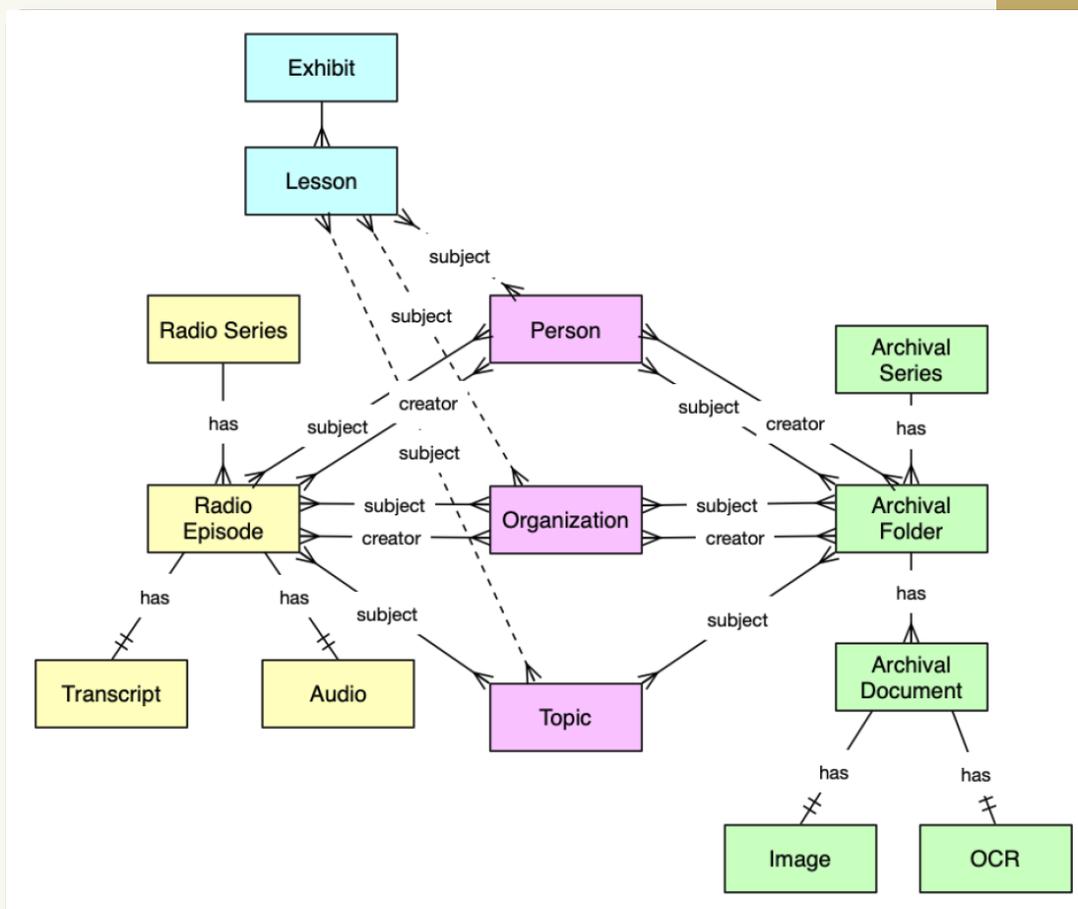
The below diagram displays all the various components of the *Unlocking the Airwaves* application, as well as how they function together as part of a process.



Appendices (continued)

[Object Model](#)

Below is the Airwaves object model, which recalls key elements and components of the project data such as radio series, radio episodes, archival series, archival folders, and authorities (Person, Organization, Topic).



Unlocking the Airwaves Object Model

[Airwaves Application Database](#)

This link gives you access to a read-only version of the Airtable database which contains not only the data feeding the *Airwaves* website (metadata about documents, programs, and entities), but also data related to existing and enhanced SNAC and Wikidata records. This database can be cloned/forked by clicking on the 'Copy base' link on the upper right, which will copy the base into the Airtable account of the users clicking on it. Users must be signed into Airtable and will need to have an Airtable Pro workspace in order to use the data as-is.

Appendices (continued)

ARCHIVAL AUTHORITIES

[Design Document: Archival Authorities Reconciliation](#)

This document describes how the Airwaves project planned for its reconciliation and enhancement of archival authorities. The exact workflow for the steps described in that document are demonstrated in two videos below.

[Reconciling Archival Authorities to Wikidata in OpenRefine](#) (video)

This video demonstrates how to use a dataset of archival authorities (in this case, people entities) to 'reconcile' (compare and match against) Wikidata, using the OpenRefine Wikidata plugin. Once the matching is completed, data can be extracted from Wikidata using OpenRefine, and then imported back into Airtable.

[Enhancing Geographic Metadata Using Airtable, OpenRefine, & Wikidata](#) (video)

This video demonstrates another piece of the archival authorities workflow by showing how the Airwaves website was enhanced using geographic metadata extracted from Wikidata using OpenRefine. This work builds off of the work described in the video above.

[Workflow for Extracting SNAC Authority Data](#)

The process for reconciling archival authorities from the Social Networks and Archival Context (SNAC) project is the same as it is for Wikidata, as described above. However, SNAC's OpenRefine plugin does not allow for data extension, which is the process of pulling data into OpenRefine using the reconciled authorities. Daniel Pitti and his team at SNAC designed an XSLT script that can be run on EAC-CPF XML files to extract data from the authorities into a tab separated values (TSV) document. That data can then be harvested/ingested into the main database (in our case, Airtable) for tracking and displaying that authority data.

Publications & Presentations

["Healing the Gap: Digital Humanities Methods for the Virtual Reunification of Split Media and Paper Collections,"](#) Peer-reviewed journal article in Digital Humanities Quarterly Volume 15.1, Spring 2021.

["On Forks and Completion: Reflections on a Five-year Digital Humanities Project,"](#) Blog post on the MITH website, July 2021.

["Workshop Report: Digital Storytelling With Audio Primary Sources,"](#) Blog Post/Report from the Radio Preservation Task Force Education Division, July 2021. (Additional links: [Presentation Video](#) | [Presentation Slide Deck](#) | [List of Digital Storytelling Resources](#))

["The NAEB Radio Collection: Document Deep South and Mississippi Waterways,"](#) Video presentation created for the 12th Orphan Film Symposium: Water, Climate, and Migration, May 26, 2020.

["Radio Collections at Scale: The Unlocking the Airwaves Project,"](#) Virtual presentation to the IIF Archives Community Group, May 12, 2020.

["Radio Collections at Scale: Unlocking the Airwaves project as a model A/V digital humanities project,"](#) Presentation at the Association of Moving Image Archivists Annual Conference, Baltimore, MD, November 2019.

["Digitizing and Reviving the Archives of the National Association of Educational Broadcasters,"](#) Presentation at the Society for Cinema & Media Studies Annual Conference, Seattle, WA, March 2019.

Project Team & Contacts

PROJECT DIRECTORS

Stephanie Sapienza (Principal Investigator, University of Maryland)
sapienza at umd dot edu

Eric Hoyt (Principal Investigator, University of Wisconsin-Madison)
ehoyt at wisc dot edu

ADVISORY TEAM

Rebecca Fraimow (Archivist, WGBH)

Megan McShea (Archivist, Smithsonian Institution)

Allison Perlman (Associate Professor, UC Irvine)

Laura Schnitker (Archivist/Curator, UMD Libraries)

Josh Shepperd (Assistant Professor, University of Colorado Boulder)

PROJECT TEAM

JJ Bersch (Production Assistant, UW-Madison)

Matthew Blessing (Administrator of the Division of Library-Archives,
Wisconsin Historical Society)

Alison Bridger (Archivist, Wisconsin Historical Society)

Emily Frazier (Graduate Research Assistant, MITH)

Paul Hedges (Archivist, Wisconsin Historical Society)

Chuck Howell (UMD Libraries)

Kirsten Keister (Designer, MITH)

Jacob Mertens (Production Assistant, UW-Madison)

Trevor Muñoz (Director, MITH)

Connor Perkins (Production Assistant, UW-Madison)

Matthew St. John (Production Assistant, UW-Madison)

Ed Summers (Developer, MITH)