# SOUND MAPPING ON THE WEB: CURRENT SOLUTIONS AND FUTURE DIRECTIONS

*Brandon Mechtley*
Arizona State University
School of Computing, Informatics, and Decision Systems Engineering

*Perry Cook*
Princeton University
Departments of Computer Science and Music

*Andreas Spanias*
Arizona State University
School of Electrical, Computer, and Energy Engineering

**ABSTRACT**

Many web-based sound maps have appeared within the last decade. In this paper, we have catalogued 94 different sound maps created since 1999 in terms of their interactive style, associated media and metadata, and content curation. Field recordings themselves provide an incredibly dense source of information and sense of place, but sound maps have yet to match visual maps in terms of ease of exploration. In this paper, we take a look at existing sound maps and discuss possible future directions for more fluid interaction styles.

## 1. INTRODUCTION

In *The Tuning of the World* [19], R. Murray Schafer wrote, "to give a totally convincing image of a soundscape would involve extraordinary skill and patience: thousands of recordings would have to be made, tens of thousands of measurements would have to be taken, and a new means of description would have to be devised." Although this possibly serves as a general caveat regarding the practicality of distilling an inherently experiential phenomenon into a finite number of recordings and measurements, it is easy to see that increased affordability and availability of portable digital audio recorders, storage media, and computational power to process vast amounts of audio data may assist in some of these challenges.

In particular, a large number of web-based sound maps have appeared within the last decade, perhaps beginning with *sound-seeker* [16] and *Soundcities* [22] in 1999-2000. Many of these tools are mash-ups with online mapping tools, having an interaction style that can be described as "point, click, and listen," where sounds are shown on a map as homogenous markers.

While this type of representation allows one to explore a location's distribution of recordings and provides a convenient method of exploring large, geo-tagged audio archives, it perhaps lacks the fluidity that allows easy exploration, as the place marks rarely indicate any features of the sounds they represent other than location, and it may be necessary to listen to several long recordings before finding one that is relevant or particularly interesting to a listener. To use a visual metaphor, it may be equivalent to clicking on individual map markers to view photographs of a city rather than the more continuous exploration allowed by satellite imagery or photo-stitching tools such as Bing Maps or Google Street View. Of course, photographic media (including satellite imagery) have the advantage of having natural spatial properties, which is more difficult to represent with point-source recordings. Nevertheless, being provided with a continuous visual flow of multiple observations makes the process of viewing maps highly interactive through the eyes alone.

Some sound archives, such as the large databases of *Radio Aporee* [14] and *Freesound.org* [25], have introduced alternative ways of listening and search, such as automatic playlist generation through random walks or folksonomy-based tag searching. In this paper, we will examine 94 different sound maps on the web[1], starting in 1999, categorizing them by their means of interaction, associated multimedia and metadata, and content curation, speculating on what future directions might exist in each of these areas. Additionally, we provide a preview for a new soundscape rendering project, generically named *Soundwalks*, which uses real-time re-synthesis and mixing of environmental audio recordings to provide a continuous scrubbing interface.

## 2. INTERACTIVE FEATURES

### 2.1. Map Embeds

Of the 94 geo-tagged sound archives listed, 80 were found to at least have some sort of *point-click-and-listen* interaction. Some of these are simply implemented as blog entries with single-marker embedded maps, presented more as a story of recording, focusing on the prose descriptions of the process as opposed to geographic relationships between sounds. However, most of these tools exist as full web sound maps, with 61 using Google Maps embeds, 1 using Bing Maps, and 5 using the Creative Commons-licensed, crowdsourced mapping service OpenStreetMap [15]. Although some of the other maps may use rather sophisticated or well-designed custom visual layers, the prevalence of map/marker mash-ups is encouraging, as it demonstrates the relative ease of developing these maps.

### 2.2. Automatic Play and Advancement

One simple factor that can hinder interaction with a sound map is the number of clicks necessary to explore it. Only 23 maps marked with *autoplay* automatically begin playing sounds upon clicking on their markers rather than having to initiate an embedded audio player, or worse, being redirected to a separate website where the sound can be played. Although fast-paced clicking is perhaps not the mindful style in which some authors

[1] An interactive listing of all 94 of these web sound maps, including those not explicitly mentioned here, is available at http://bmechtley.me/websoundmaps.html.

wish their listener to engage, being able to smoothly shift from one sound to another can make the resultant soundscape much more seamless and allow for quick comparison of acoustic properties between locations.

Another possibility for automatic play, referred to as *autoplay next*, is automatic advancement to other sounds on completion, featured in maps such as the *Montréal Sound Map* [23], *Favourite Sounds* [6], the *Brussels SoundMap* [9], and *Cartophonies* [5]. This feature is most useful for maps that include sound sequences and paths, as will be described in Section 2.3, but even allowing for random walks between neighbouring recordings, such as in *Radio Aporee*, allows a listener to sit back and actually listen, using the map as a visual reference, rather than dedicating significant attention to mouse movements and interaction with audio widgets.

### 2.3. Sequences and Paths

6 of the maps listed allow temporal sequencing of sounds. Sound sequences are especially interesting for the recording of the soundwalk. Some field recordings taken during soundwalks involve several segregated sound files. Being able to both initiate a sequence of sounds and have the player automatically advance to the next in the sequence allows the author to relate a sonic narrative. On some sound maps, such as *Listen to Africa* [27], knowledge of the path the recordists travelled and inspection of the geographic distribution of the placemarks is enough, but in dense archives such as *Freesound.org* and *Radio Aporee* (which provides this feature, along with 6 other maps), it can sometimes be necessary to organize a sequence separately.

An obvious extension of the *sequences* pattern is to visually display connections between the sounds through *gps traces*. In the simplest case, since most audio archives record the time and date of recording, this can be as simple as drawing line segments between sounds of the same recording session or soundwalk. In more detailed settings, continuous GPS traces from external devices can be used, which can easily be obtained from most smartphones. In retelling the story of a soundwalk, in particular, these GPS traces convey important information about the walk, such as the moment-to-moment decisions made in improvisational exploration, paths left silent, or prepared composition. These traces are especially important in the presence of lengthy (e.g. several hour) non-stationary recordings that are not segmented. As of yet, no web sound map seems to address the issue of conflating an entire, continuous mobile field recording onto a single point.

### 2.4. Mixing

Dynamic *mixing* of multiple field recordings is a particularly interesting pattern supported by 5 of the maps, namely the *Cinco Cidades Soundmap* [7], *Tactical Sound Garden* [21], *Favourite Sounds*, *soundingD SoundMap* [12], and *Radio Aporee*. Most of these tools allow for manual selection of several recordings to mix, which potentially allows the map itself to be a live performance or compositional tool, likely of interest to phonographers. One potential artefact (or affordance, depending on one's outlook) of these strategies is that directly mixing multiple soundscapes can result in a cacophony of overlapping acoustic streams that cannot be segregated psychoacoustically [3]. Using Schafer's categorization, this can result in obfuscation of keynote sounds, not to mention interrupted signals and soundmarks.

### 2.5. Mobile Applications and Live Streaming

Two additional patterns, supported by only a few platforms, include the presence of *mobile* applications and the mapping of live audio streams. *Audioboo* [2], a popular audio-blogging site, initially contained a sound map of recordings as its landing page, though they have since focused on a more social networking style, where the geographic metadata has unfortunately been relegated to the page view for individual sounds. The enterprise *SoundCloud* [24] can also provide GPS metadata when recordings are uploaded from its mobile application, but as of yet, besides custom pages that position them, it has not explored audio mapping as an official facet of its business. Two other non-commercial tools, however, do provide mobile recording, including the soundscape affect research project *Sound Around You* [13] and *Radio Aporee's* recent mobile version.

Live streaming is an interesting modality that has existed for some time, and though only 3 of the listed platforms [1, 10, 18], aggregate multiple streams on a single map, a live stream does provide an excellent sense of a location's temporal activity patterns. *LIDO*, in particular, has an interesting real-time spectrogram of its partnered live streams. Many other solitary live audio streams (mostly focused on bioacoustics applications) exist and could potentially be aggregated on a map.

## 3. MULTIMEDIA AND METADATA

### 3.1. Other forms of media

While reduced or musical listening approaches to the soundwalk may deemphasize the context of a sound stream, providing additional media related to a recording can be quite helpful in recording its experiential context. 16 of the maps provide one or more images alongside the recordings, and 3 of those also provide video content. These additions are especially common amongst field recording blogs. *Xeno-canto* [28] and *LIDO* and *Freesound.org* both provide spectrograms, which could potentially be interesting data to integrate more seamlessly into a map. The various sound maps on the *London Sound Survey* also provide additional layers, such as historical maps.

### 3.2. Text description

14 of the maps provide each sound with folksonomy-based single-term tags, some of which allow browsing or searching by these tags. 28 of the maps provide segregated categories or taxonomies, some in particular being inspired by the taxonomy first proposed by Schafer. More generally, 70 maps provide some description of the sounds being played, though often these can be limited to descriptive filenames. 12 maps provide all three descriptors. Social media repositories tend to trend toward folksonomy-based tagging, but

explicit categorization can quite useful, especially when a comparative study is desired. While not present on all maps, this type of metadata is one step away from providing a more investigative experience for listeners. Content-based search, such as techniques used in music information retrieval or speech analysis are also obvious candidates, though as of yet unexplored in these maps.

## 4. CONTENT CURATION

### 4.1. Community involvement

The 94 sound maps also vary drastically in how they curate content. 69 of the maps are either now closed or limited to the contributions of the authors. The remaining 22, however, allow some sort of community contribution, varying from automatic submission interfaces to explicit request via email. The crowdsourced data brings up obvious issues of licensing restrictions. It is fortunate, then, that the larger projects, *Radio Aporee*, *Freesound.org*, and *SoundCloud* allow users to choose between various licensing schemes, including Creative Commons options. *Sound Around You* has an interesting approach from a research standpoint, where users are not only able to upload their sounds from mobile devices, but are also able to comment on the soundscapes' affective properties.

### 4.2. Other types of audio

While field recordings largely (or solely) make up 76 of the listed projects, there are some sound maps that focus explicitly on music or speech, such as some of the mapping projects by The British Library [4]. *Audioboo* likely has a similar focus on speech, with its branding as an audio blogging tool. Additionally, a few of the soundmaps focus at least in part on historical data, including the soundmaps of the *London Sound Survey* and efforts by The British Library, the latter of which include recordings from historical archives.

## 5. *SOUNDWALKS:* CONTINUOUS PLAYBACK

### 5.1. Application architecture

The demonstrated web application, *Soundwalks*, attempts to elaborate on some of the topics mentioned in the previous sections. Most importantly, it addresses issues related to uninterrupted exploration of the soundscape and visualization of the soundwalk path. In its current form, *Soundwalks* is a web application intended for personal archival and reflection upon acoustic journeys. *Soundwalks* works as an HTML5 application that uses the WebSocket API to communicate to a local server running "pyo," a Python-based modular synthesis library. Optionally, the server can also route messages from the browser as OSC to be used by other synthesis applications or interfaces such as PureData, Max/MSP, ChucK, etc.

### 5.2. Interaction through scrubbing

The user interface shown in Figure 1 is sparse, showing a list of soundwalks with associated titles and

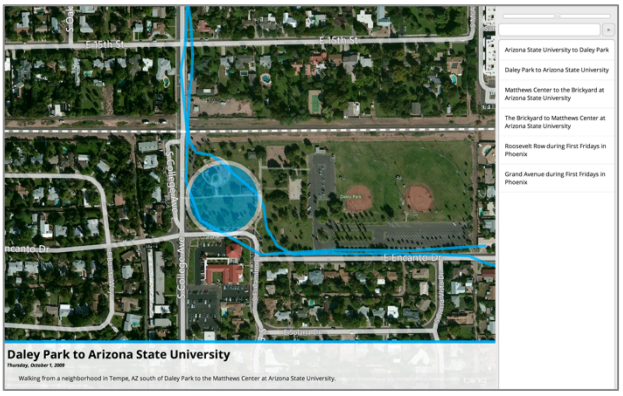


**Figure 1**: An example of *Soundwalk's* synthesis interface. The cursor is situated between two paths.

descriptions that move the map to their coordinates when clicked. Every full GPS path, imported via GPX from a mobile application or dedicated GPS receiver, is shown. As the user drags a cursor across the map, regardless of whether or not a path is underlying, a continuous soundscape is played. This soundscape is a dynamic mixing of surrounding recordings that uses a two-stream model to blend acoustic environments. This two stream-model attempts to resynthesize sounds in a method inspired by auditory streaming phenomena [3] and other figure/ground-based classifications of the auditory scene. Additionally, when mixing textures, it avoids problems of overlapping multiple background textures and interrupting foreground sonic events.

### 5.3. Multi-stream synthesis

Upon inclusion into the map, each recording is first segmented into separate audio events using a probabilistic event detection model described in [26]. While this segmentation has a number of phenomenological implications, relying on the debatable definition of "sonic event," it assists in separating short-term tonal or transient events from what might otherwise be considered an acoustic ground.

The first stream, or ground, is a resynthesized audio texture resulting from a parametric blend of the remaining background segments clips from each soundwalk within a defined radius of the user's cursor. This mix is constructed through a variation of the wavelet tree learning technique in [8], which attempts create iterations on audio textures while retaining their structural properties as opposed to more common granular synthesis clouds or concatenative synthesis (for a recent summary of audio texture synthesis techniques, see [20].) The soundwalks closest to the cursor are favoured in this re-synthesis process. The second stream, or figure, is constructed by a stochastic triggering of the extracted sound events, favouring those closest to the cursor, and attempting to match the average event density from the surrounding recordings.

### 5.4. Implications of the synthesis model

Some benefits of a real-time mixing and scrubbing interface include the explorative navigation more characteristic of visual maps, where the same type of

instant navigational decisions can be made, as the listener has constant, immediate feedback of his actions. Additionally, explicit inclusion of continuous GPS paths for each mobile recording allows viewers to better understand the geographic context and stream of decisions made by the recordist in the moment.

However, it should be noted that this method is by no means meant to be an end-all-be-all solution to acoustic mapping on the web. As has been detailed in the survey of existing sound maps, there are a wide variety of sonic cartography projects, each with their own theoretical backings, artistic statements or intent, or preferred listening style.

## 6. CONCLUSIONS AND FUTURE WORK

This paper serves as an initial description of the state of sound mapping on the web, what appears to be a burgeoning medium for reflection, expression, and scientific exploration. It is clear from past and present projects that there is great variety in the intent of every map, but much is still to be explored in terms of *expressing* that variety through their demonstration.

Through our own works with the *Soundwalks* project, there are many future directions that are being explored in terms of how to best blend between multiple environmental recordings and how to convey a more informative and engaging experience of a soundscape. Specifically, work in automated annotation and retrieval of environmental audio [11] can assist in providing visual displays of the geographic distribution of associated metadata, such as tags or acoustic features.

## 7. REFERENCES


[1] André, M., van der Schaar, M., Zaugg, S., Houégnigan, L., Sánchez, A. M., and Castell, J. V. "Listening to the Deep Ocean Environment," listentothedeep.net, 2011.

[2] Audioboo, Ltd. "Audioboo," audioboo.com, 2011.

[3] Bregman, A. S. *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT Press, Cambridge, MA, 1994.

[4] The British Library Board. "British Library Sound," sounds.bl.uk, 2004.

[5] Cheikoff, G., Acquier, F., Moisans, J., Laroche, S., and Bérubé, G. "Cartophonies," cartophonies.fr, 2012.

[6] Cusack, P., Currie, N., and Edwards, K. "Favourite Sounds," favouritesounds.org, 2011.

[7] Dant, A., Davis, T., and Gunn, D. "Cinco Cidades Soundmap," cincocidades.com, 2007.

[8] Dubnov, S., El-Yaniv, R., Lischinski, D., and Werman, M. "Synthesis of Sound Textures by Learning and Resampling Wavelet Trees." *IEEE Computer Graphics and Applicotions*, 2002.

[9] Harry + LVH. "Brussels SoundMap," bna-bbot.be/brusselssoundmap, 2011.

[10] Locus Sonus. "Locustream Map," locus onus.org/soundmap/, 2006.

[11] Mechtley, B., Cook, P., and Spanias, A. "Shortest Path Techniques for Annotation and Retrieval of Environmental Sounds," in *Proceedings of the International Conference on Music Retrieval*, Porto, pp. 541-546, 2012.

[12] Minard, R. and Ludger, H. "soundingD SoundMap," map.sounding-d.de/map.

[13] Mydlarz, C. "Sound Around You," soundaroundyou.com, 2009.

[14] Noll, U. "Radio Aporee," aporee.org/maps, 2006.

[15] OpenStreetMap contrib., "OpenStreetMap," openstreetmap.org, 2004.

[16] Polli, A. and Feng, S. S. "Sound-Seeker," soundseeker.org, 1999.

[17] Rawes, I. "London Sound Survey," sound survey.org.uk, 2011.

[18] Salish Sea Hydrophone Network. "Orcasound.net," orcasound.net, 2011.

[19] Schafer, R. M. *The Soundscape: Our Sonic Environment and the Tuning of the World*. MIT Press, Cambridge, MA, 1994.

[20] Schwarz, D. "State of the Art in Sound Texture Synthesis," in *Proceedings of the 14th International Conference on Digital Audio Effects (DAFx-11)*, Paris, pp. 221-231, 2011.

[21] Shepard, M., Murphy, F., Diesel, B., Flynt, A., Thomas, A., Modi, V., and Krishnamurthy, A. "Tactical Sound Garden," tacticalsound garden.net, 2007.

[22] Stanza. "Soundcities," soundcities.com, 2000.

[23] Stein, M. and Stein, J. "Montréal Sound Map," montrealsoundmap.com, 2008.

[24] Soundcloud Ltd. "SoundCloud," sound cloud.com, 2007.

[25] UPF Music Technology Group, "Freesound," freesound.org, 2005.

[26] Wichern, G., Xue, J., Thornburg, H., Mechtley, B., and Spanias, A. "Segmentation, indexing, and retrieval for environmental and natural sounds." *IEEE Transactions on Audio, Speech, and Language Processing*, 18(3):688-707, 2010.

[27] Williams, H. and Summers, R. "Listen to Africa," listentoafrica.com/map, 2009.

[28] Xeno-canto Foundation. "xeno-canto," xeno canto.org, 2005.