Audio- & Visual-Equivalence Guidelines for Digital Signage

Last updated by **Technology Innovation** on **June 7, 2022**.

Background

The proliferation of digital signage on the MBTA system, generally, and the creation of a "Screens" program within the MBTA's Technology Innovation Department (TID), has created both a challenge and an opportunity:

> *How do we make sure that MBTA riders who are blind or low-vision get an equitable experience when we make information available on digital signage?*

And because the MBTA makes audio announcements in a variety of contexts—in our rapid-transit stations, onboard our trains and buses, etc.—this has, in turn, yielded a related question:

> *How do we make sure that MBTA riders who are deaf or hard-of-hearing get an equitable experience when we make information available primarily via audio announcements?*

The majority of this document—and of our work to date—is in responses to the first question, about audio-equivalence for blind or low-vision riders. This is in large part because the motivating problem is the expansion of visual information via digital signage. There are scenarios in which audio announcements are the default, and visual-equivalence is far behind, but these are less common (see "Cases Studies" below).

Purpose

The purpose of this document is:

- To articulate a policy decision that we hold ourselves accountable to;
- To document the decisions we've made to date, and the thinking behind them, which serve as guidelines for future projects led both internally and by non-MBTA third parties (e.g., consultants, developers);
- To share what we have and have not learned about audio- and visual-equivalence, in order to keep learning from riders, advocates, peers, and industry partners.

Scope

The scope of this document is the *how* of audio- and visual-equivalence in the context of digital signage[1]—and, primarily, digital signage whose purpose is to deliver real-time information (i.e., live-updating information about the system or the service) to riders.

It does not cover:

- *What* information should be made available in audio and visual formats in any given situation
- Technical or design guidelines for audio and visual information delivery (e.g., type size, audio decibel level, etc.)
- Audio-equivalence in the context of physical (i.e., non-digital) signage created by the MBTA or non-MBTA third-parties (including both Wayfinding signage and temporary signage)
- Screens on fare vending machines (FVMs)

Policy Statement

Digital signage projects must make information available in audio and visual formats.

General Approach

We expect that each digital signage project or deployment will require context-specific decisions and tradeoffs. This document therefore proposes a four-part framework to support clear, consistent decision-making.

Before getting to the framework, these are general guidelines to provide some context for the framework itself:

- **We do not have a universal, silver-bullet technology solution to this challenge**. We wish we did. Most importantly, this means that we are open to learning more about how other transit agencies and advertising companies are addressing the question.
- **We are not striving for precise, one-for-one equivalence between audio and visual information**. There might be good reasons to edit, or tailor, either one in service of clarity or concision. These include technology limitations, the experiential differences of audio vs. visual information, the relative importance of certain information in specific transit contexts, and more. (See “Case Studies/Examples” below for examples of decisions and trade-offs that we’ve made.)
- **Riders should know where to expect equivalent information**. In other words, trying to make all information available everywhere leads to a deterioration in the quality of information delivery. Creating a clear expectation of where and how to get information

[1] In the transit industry, digital signage—or digital screens—are often referred to as “variable message systems,” “customer information displays,” “countdown clocks,” and sometimes interchangeably with “passenger information systems.”

from the MBTA is an important part of solving for equivalence. (See “Case Studies/Examples” below for examples.)

- **Our data suggest that button-activated audio is an important part of an overall solution to audio-equivalence**. Two years ago, we thought that the industry-standard button-activated audio was probably not very useful or widely-used by riders. But a two month-long data investigation led us to change our mind: we didn’t notice any use of button-activated audio by blind riders, but we noticed it used often by older adults and by station officials tasked with helping riders navigate the system. (“Often” is subjective. The data showed 10 “useful” activations per screen per month.) So we now include button-activated audio in all screen procurements, at least for screens at stops and stations whose role is delivering MBTA service information.

Framework

Audio-Equivalence

For the purposes of audio-equivalence, this document proposes a four-part framework, which borrows from the Federal Highway Administration’s Manual for Uniform Traffic Control Devices (MUTCD).

The context for this framework is a scenario in which visual information is being made available on digital signage.

Stage	Solution Space
How does a blind or low-vision rider know that there is visual information in the current environment?	Some ways to address this: • Continuous public address (no knowledge required) • Screen emits a locator tone • Proximity- or motion-based audio actuation • Button (requires discovery)
How does a rider actuate the receipt of that information?	There is a spectrum from: • No action is required (e.g., a continuous public address system) • Passive action is required (e.g., motion-detection-based actuation of an audio feed) • Active actuation is required (e.g., a button or headphone jack)
How does a rider receive the information?	Some ways to address this: • A public address system • Parametric (i.e., directional) speakers • Personal headphones • Through a personal smartphone

How is equivalence maintained?	Broadly, there are two approaches: • Strict equivalence with visual information; OR • Context-specific adjustments that break one-for-one equivalence See “Case Studies” below.

Visual-Equivalence

We propose using the same four-part framework as above, acknowledging that that might not be the best or more useful approach. (This is itself an area for improvement.)

The context for this framework is a scenario in which audio information is being made available primarily via a public address system.

Stage	Solution Space
How does a deaf rider know that there is audio information in the current environment?	There should be some visual indication on a screen in the given environment *that the rider is likely to notice*
How does a rider actuate the receipt of that information?	[we don’t have examples of actuation being required]
How does a rider receive the information?	Visual information on a screen
How is equivalence maintained?	Broadly, there are two approaches: • Strict equivalence with audio information; OR • Context-specific adjustments that break one-for-one equivalence See “Case Studies” below.

Case Studies/Examples

Public Address & Electronic Signage System (PA/ESS)

This is the primary and most ubiquitous form of digital signage on the MBTA system: 450+ overhead-mounted “countdown clocks” in every rapid transit station and 10+ major busways, with speakers.

How we've applied the framework

Stage	Solution Space
How does a [blind/deaf] rider know that there is [audio/visual] information in the current environment?	Audio: continuous public address Visual: overhead, monochrome LED screens
How does a rider actuate the receipt of that information?	No actuation required
How does a rider receive the information?	Speakers and screens installed throughout stations & busways
How is equivalence maintained?	The same software powers both audio and visual information. Other than abbreviations of station names, punctuation, etc., strict equivalence is maintained for all types of messages: • Train/bus arrivals • Service alerts • Ad-hoc messages (e.g., "No trains; use shuttle bus) posted by the Operations Control Center • PSAs (e.g., "Face coverings are required…") • Pre-written emergency messages The only exception is the "live public address system" (livePA), which allows the OCC to make immediate, audio announcements from a microphone into a station. The same system does allow the OCC to make these announcements via pre-written messages and ad-hoc, written messages, either of which would appear in stations in both audio & visual formats.

Challenges & Open Questions

- Overhead-mounted, low-resolution monochrome LED screens are extremely limited in what information they can display at an accessible type size. The downside is that it limits/constrains our ability to share complex service information. The upside is that this limitation forces the MBTA to be extremely concise, and to avoid the temptation to use the PA/ESS system for non-essential information.
- We do not have good, research-based guidelines for things like:
    - How often the public address system should play
    - How to calibrate hardware and software to balance the competing needs of "audio that can be heard everywhere" and "audio that isn't abrasive to noise-sensitive riders"

Real-time dedicated screens in subway stations & busways

These are mostly newer screens, mounted at ground-level in rapid-transit stations—mostly in lobby areas, and occasionally in busways, too.

How we've applied the framework

Stage	Solution Space
How does a [blind/deaf] rider know that there is [audio/visual] information in the current environment?	Requires the rider to know of the screen's location & to locate a button
How does a rider actuate the receipt of that information?	Button-activation
How does a rider receive the information?	Speakers built into the screen enclosure or button
How is equivalence maintained?	We do not maintain strict audio-equivalence the way a screen reader does. Audio messages are shortened in various ways to

	be as concise and contextually-helpful as possible. Two examples: • Visual and audio from a screen in a busway • Visual and audio from a screen in a pre-fare lobby

Challenges & Open Questions

- We departed from the screen-reader-like approach so that we could tailor audio messages to a particular context, and use an MBTA-specific lexicon, rather than a default robo-voice. But that means that we can't offer riders the ability to "skip" to the next section, the way a screen reader allows. We don't know of hardware solutions provided by the digital signage industry that would enable all of these features at the same time.

Real-time dedicated screens INSIDE bus shelters

How we've applied the framework

Stage	Solution Space
How does a [blind/deaf] rider know that there is [audio/visual] information in the current environment?	Audio: continuous audio, on a timer
How does a rider actuate the receipt of that information?	No action required
How does a rider receive the information?	Speakers built-in to the screen enclosure
How is equivalence maintained?	We do not maintain strict audio-equivalence the way a screen reader would. Audio messages are shortened in various ways to be as concise as possible, especially because of it being continuous. An example: • Visual and audio from a bus stop

Challenges & Open Questions

- Continuous audio increases the need to make audio messages as concise as possible. Audio messages currently only contain bus arrivals. They'll soon include high-priority service alerts, too. But if they also included low-priority service alerts, PSAs from the MBTA, and all other visual content displayed on the screens, this would result in both a dilution of quality (i.e., bus arrivals would play *less* often) and noise pollution (in that audio would be playing almost all of the time). We don't have good research from blind riders on how to strike this balance.

Solar-powered electronic ink (E Ink) signs at bus stops & surface stations

We will soon have ~40 double-stack, 13'' E Ink screens at bus stops and surface light-rail stops, dedicated to the display of real-time information. We are in the process of scaling up to 100+ screens.

How we've applied the framework

Stage	Solution Space
How does a [blind/deaf] rider know that there is [audio/visual] information in the current environment?	Audio: in the pilot phase, we didn't provide any audio-equivalence; in the scale-up, we will add a button, which will require riders to know of the screen's location & to locate a button

How does a rider actuate the receipt of that information?	Button-activation
How does a rider receive the information?	Speakers built into the screen enclosure or button
How is equivalence maintained?	We will not maintain strict audio-equivalence the way a screen reader would. Audio messages will be shortened in various ways to be as concise and contextually-helpful as possible.

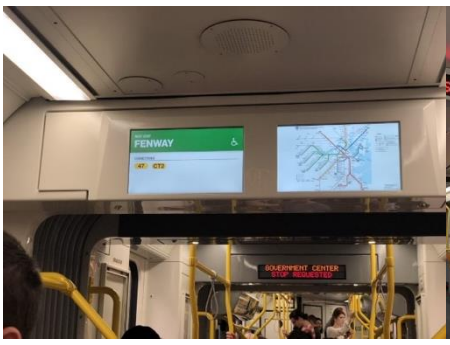
Challenges & Open Questions

- We do not know whether or how the frequency of audio activations will affect battery life and, as a result, screen uptime in low-sunlight conditions (e.g., winter, consecutive cloudy days, etc.).

LCD screens on newer rapid-transit vehicles

The Green Line Type 9s have two, overhead-mounted, landscape-orientation LCD screens in each car. These screens can do just one thing: display a single, stop announcement image file for each station.

The new Red Line and Orange Line vehicles have three, landscape-orientation LCD screens in each car. These screens can display image or video files in accordance with any automated stop announcement, or any ad-hoc message that is pre-loaded onto the vehicles.



How we've applied the framework

Stage	Solution Space
How does a [blind/deaf] rider know that there is [audio/visual] information in the current environment?	The LCD screens are in the same environment as other audio and visual communication channels.

	Audio: automatic stop announcements OR ad-hoc/emergency audio messages from the Motorperson Visual: overhead, monochrome LED screens for automatic message; no visual-equivalence for ad-hoc or emergency messages from the Motorperson
How does a rider actuate the receipt of that information?	No action required
How does a rider receive the information?	Audio: speakers located throughout the train car Visual: overhead, monochrome LED screens
How is equivalence maintained?	We do not maintain strict equivalence. We allow the LCD screens to contain *additional* or *supplementary* stop announcement information, for which we don't provide any audio equivalence. Specifically, we display the list of connecting bus routes, Commuter Rail lines, and official, 3rd party connecting services (e.g., Amtrak), but omit this information—for reasons of length—from the automatic audio stop announcements. • Visual from the Green Line Type 9s • Visual from the new Orange Line vehicles

Challenges & Open Questions

We maintain strict equivalence between audio and visual (on the LED screens) automatic stop announcements.

But we do not know how to provide:

- Audio-equivalence for the additional, or supplementary, stop announcement information on LCD screens. We could imagine that, in future, this could include real-time information like the departure time of bus connections at the current stop.
- Visual-equivalence when Motorpersons make ad-hoc or emergency announcements over the PA system.
- *Good* visual-equivalence for messages on the Type 9 vehicles that only display on the overhead, monochrome LED screens. E.g., when a train is set to run express and this

information is only conveyed on the LED screen, we know that riders who are deaf or hard-of-hearing are very unlikely to see this information, and report feeling trapped on the train.

Open Questions

We intend for this to be a living document that is updated as we learn more and come to new decisions.

The following are some specific areas that we hope to learn more about:

- How can we better **help blind and low-vision riders know of visual information in the environment in the first place**? Are locator tones a useful concept? Or would these contribute to ambient noise pollution? What about QR codes like these?
- Are there off-the-shelf **hardware solutions that are more feature-rich than buttons as audio-activation methods**? Ones that might allow for contactless activation? What else exists in the digital signage industry?
- We are likely 5-7 years away from retiring vehicle models that have only very basic LED screens for visual information, and on which we depend on Motorpersons to make manual, audio-only announcements for any ad-hoc information (e.g., trains skipping a stop, train going express, elevator out-of-service). What can we do in the meantime to **improve visual-equivalence of on-vehicle announcements** in the meantime?
- Acknowledging that our work has, so far, been more focused on audio-equivalence than visual-equivalence, what would **a better framework, and better version of this document be, on the topic of visual-equivalence** for riders who are deaf or hard-of-hearing?