

Public Radio Podcast Measurement Guidelines

Link to this document: bit.ly/podcastguidelines

Contents:

[Introduction](#)

[Contributors](#)

[Goals](#)

[Definitions](#)

[Technical Guidelines](#)

[Measurement Choices](#)

[Details of Internal Measurement](#)

[Counting Unique Downloaders](#)

[Counting Unique Downloads](#)

[Audits](#)

[Further Work](#)

[Contacts and Feedback](#)

Introduction

Measurement of podcast usage is, at best, the Wild West. There has been no standard or even informal consensus around how to count downloads, listeners, or time spent listening. Given the attention to and growth of podcasts and on-demand audio in general,

inconsistent approaches to podcast measurement have hampered the ability of this medium to reach its full competitive potential as an alternative to broadcast and other sponsorship media with well-defined and widely adopted audience measurement and performance standards. Uniform and fully understood on-demand audio data is essential to fully realize the business opportunities of this medium for podcast producers, sponsors and other stakeholders.

Beginning in the spring of 2015, the major producers of public radio podcasts have been convening to discuss podcast measurement and develop our own guidelines to establish greater consistency and better accuracy for measuring on-demand audio performance, and to foster a uniform understanding of data that can be used by all stakeholders to make informed decisions and comparisons.

This document is the result of those meetings. These guidelines are not intended to operate as a full technical standard per se, but rather overall principles and public radio's technical guidelines for measuring podcast usage. As representatives of mission-driven organizations that increasingly rely on podcast sponsorship as a significant part of their fundraising efforts, the members of the working group believe these guidelines are an important first step toward creating a more robust, transparent sponsorship medium for the mutual benefit of public media outlets and their underwriters. Although designed primarily as open and voluntary guidelines for public radio podcast producers, the guidelines are also open to voluntary adoption by other podcast producers, providers of third-party tools, vendors, sponsors, advertisers, agencies, and standards bodies. Indeed, the guidelines will be most effective in achieving their goals if they are voluntarily adopted by numerous participants in the on-demand audio ecosystem.

This document provides guidelines to assist organizations in creating consistent, reliable data for more efficient evaluation and analysis of podcasts and other on-demand audio services. However, these guidelines do *not* suggest ways to monetize podcasts or on-demand audio. Decisions about pricing and monetization are outside the scope of these guidelines and all such decisions must be made independently and unilaterally by each individual entity.

Contributors

The following people and organizations participated in the creation of these guidelines.

Stephen Haptonstahl, NPR *

Dan Jeselsohn, WNYC *

Steve Mulder, NPR *

Eric Richardson, SCPR *

Kyle Wesloh, MPR/APM *

Leng Caloh, KPBS

Tiffany Campbell, WBUR

Morgan Church, PRI

Joe DeCeault, WBEZ

Marissa DeMartini, WBEZ

Tom Interrante, WXPB

Bill Irwin, WNYC

Nathan John, KPBS

Andrew Kuklewicz, PRX

Alex Lee, CBC

Chris Lewis, WAMU

Kai Lucas-Baradan, PRI

Joel Meyer, WBEZ

Bryan Moffett, NPM

Lauren Morris, WNYC

Manuel Ojeda, KQED

Stefan Olson, MPR/APM

Mathilde Piard, NPR

André Turcotte, CBC

Colleen Wilson, KQED

* Steering committee

Goals

The goals of this initiative include the following:

- **Fostering better intelligence for public radio podcast producers and industry analysts.** Without a common measurement approach, there is simply no way to accurately compare podcasts from different organizations or explore industry trends. These guidelines provide an apples-to-apples baseline that enables better intelligence on what's working.
- **Creating consistency and credibility for sponsors.** Currently there is no industry-wide or consistent methodology that engenders the kind of sponsor confidence needed to maximize the potential of on-demand audio as a sponsorship medium. Sponsors deserve transparency. Establishing a standard measurement methodology will provide our industry with credibility and will attract new, more diverse sponsors.
- **Setting precedent.** On-demand audio does not benefit from the same degree of standardization that video and other mature media enjoy. These guidelines are setting precedent for the future of on-demand audio by establishing explicit definitions, metrics and a framework for how measurement will evolve.
- **Leveraging industry know-how to promote standards consistent with the needs of public media.** By leveraging the unique technical knowledge and experience acquired by public media organizations as leaders in podcasting, these guidelines will help ensure that open standards for the measurement of on-demand audio services will continue to evolve in ways that harmonize with the needs and operations of public media. If public media stakeholders do not work together now to define standards for the measurement of on-demand audio, others outside of our industry will do so and may establish parameters that are incompatible with public media operating environments.

Definitions

“Podcast” is a slippery label, once defined by its technical delivery platform but increasingly used to describe an entire class of audio content. These guidelines are applicable for all on-demand audio, and adopt the following terminology and definitions for clearer discussions moving forward.

On-demand audio is broad. It refers to any digital audio downloaded by request and not listened to via a live stream. (A live stream refers to many people tuning in simultaneously and hearing the same thing.) The downloading can occur via any platform or file transfer protocol, including progressive downloads on a web page, downloads to an app, or plays in iTunes. On-demand audio includes full shows/podcasts, segments from shows/podcasts, and standalone audio stories or clips.

Podcasts are a subset of on-demand audio. They consist of recurring shows or audio content collections. Measurement of downloads should include any form of on-demand, digital listening to that podcast, regardless of platform and inclusive of full episode downloads and downloads of segments of an episode. Often this is limited to audio files downloaded because they were enclosures in an RSS feed but may also include things like download links on a Web page or plays of an episode via a Web-based player.

Basic measurement of shows/podcasts includes all downloads, but some organizations might also need to generate a separate measure for sponsorship purposes that includes a subset of all downloads.

Technical Guidelines

The guidelines that follow are primarily for organizations who are parsing and generating metrics from raw log files. Many organizations use third-party tools or vendors for podcast measurement. Ideally, these vendors will also adopt the following guidelines, so that everyone gets consistent data.

When these guidelines use the phrase “will be” in reference to a particular action (e.g., “will be used/included/filtered/reported/counted, etc.”), this language is intended to indicate

what an organization must do to in order to say that it complies with these guidelines. Outside of these parameters, participating organizations are free to choose their methods of implementation.

Measurement Choices

Some organizations rely on outside vendors for their metrics; some generate their metrics internally. Clearly, if metrics are available from only one of these (either external or internal metrics) then those metrics will be used. If metrics are available from both, it is strongly encouraged that the primary source of metrics should be those that adhere closest to the guidelines outlined in this document. Internal and external metrics can be used in concert; for example, a program might release overall download numbers reported by an external vendor but report downloads by Nielsen Designated Market Area based on internal analysis because that breakdown is not available from that external vendor.

Details of Internal Measurement

Internal measurement is assumed to be based on the analysis of download logs. These guidelines assume that for each time a podcast app communicates with the server (an http request) one line is written to the logs. For all metrics, a series of filters should be applied. Such filters do not guarantee that downloaders listened to the files that remain, but they are designed to remove traffic that is known not to result in actual listening.

Requests via GET will be retained and other types of requests (such as HEAD) will be filtered. This removes requests for things like “are you there?” and limits it to requests to actually download the file.

Requests with http response codes 200 and 206 will be retained; all others will be filtered out. This eliminates all non-relevant server responses, such as 302 redirects (“no, the file you want is over there”) and limits it to requests for either the whole file (200) or part of the file (206). For 206 requests, the byte range must either be “-” or “0-*” (a hyphen or start with zero). For those organizations parsing Akamai logs, 000 responses will be retained where the byte range is either “-” or “0-*”, as with 206. The byte range filter for 206 and 000 ensures that a file is counted only once when it is downloaded in multiple parts.

Requests with user agents identifying known bots will be filtered. Some bots (like search engines) identify themselves, so one can filter out their automated requests.

Each participating organization will make a good faith effort to identify downloads that don't appear to represent actual potential listens. For example, if you get 10,000 downloads of one show to one IP address in a day, it is unlikely that those represent downloads that will be heard by listeners. Where possible, these downloads will be identified and filtered out. Specific thresholds may vary by show and organization, and is left to the organization to determine what thresholds should be used in the context of its operations, consistent with the overall goal of maximizing accuracy.

Requests from IP addresses used by the organization, including those of the CDN or other company serving the files, will be filtered out. This removes test downloads and downloads by organization employees while at work.

Requests that are served less than 200,000 bytes will be excluded. Some services often make requests for small parts of a file without delivering those to a listener. Experiments have shown that this threshold can eliminate most of this traffic without filtering out small size episodes.

Counting Unique Downloaders

It's difficult to count accurately the number of downloaders: no unique ID is transmitted when requesting a podcast file; multiple downloaders can use a single IP address (such as when they are on a shared private network); one downloader can have multiple IP addresses (such as when changing cellular towers). Each downloader does transmit a user agent description which varies by software and sometimes by hardware used. The combination of IP address and user agent provide something closer to a unique identifier for a device, which is itself an approximation of a unique identifier for a downloader.

Where the user agent of the requesting client is available, this will be a count of the unique combinations of IP address and user agent for the period reported. Otherwise, this will be a count of unique IP addresses for the period reported.

Counting Unique Downloaders

downloaders (over a time period like a week) = GET +
200 or (206 + (byte range 0-* or -)) +
no bots or filtered IPs +
> 200,000 bytes +
unique (IP + user agent) or unique (IP + user agent + show)

If downloaders are reported weekly, this is not the sum of daily unique downloaders but rather values (of either user agent-IP address or just IP address) unique across the entire week reported. This will be reported with language such as "There were 2,325 unique downloaders for this show for this week." Similarly, uniques for an organization will be uniques that are distinct across shows, not the sum of uniques for each show. This will be reported with language such as "There were 10,253 unique downloaders for this station for this week." It is appropriate to report the average of uniques across time periods, such as "There were 34,121 average downloaders per day in June."

Counting Unique Downloads

Logs often break a single download into more than one line indicating that more than one request was involved in the download. To get an accurate count of downloads, these lines must be merged based on some criteria. A unique download is defined by a unique combination of the ip address, user agent, file path/name, and the date of the download. A result of this is that two downloads by the same user (IP address-user agent) of the same file in the same day will be considered one download, but two downloads by the same user of the same file on two different days will be considered two downloads.

If the requested byte range and the size of the file are available, requests where the beginning of the byte range is the size of the file or greater will be filtered out.

Counting Unique Downloads

downloads = GET +
200 or (206 + (byte range 0-* or -)) +
no bots or filtered IPs +
> 200,000 bytes +
unique (IP + user agent + file + year + month + day) +
byte range start < file size

Audits

Auditing is encouraged to ensure compliance with the guidelines described above if the participating organization's on-demand audio data is anticipated to be used by third parties. Regular audits are recommended for both the processing of the data, as well as for the data itself (i.e., actual physical collection testing).

Further Work

The guidelines presented in this document have the greatest impact when adopted by the greatest number of organizations. Accordingly, this group intends to work together to raise awareness of these guidelines by other podcast producers, third-party tools, vendors, agencies, and standards bodies. Given the leading position of public radio podcasts as a source of podcast listening across the country, this group believes these public radio guidelines could become the foundation for a true industry standard for all podcast measurement.

The group also intends to continue to collaborate on other aspects of podcast measurement, for example:

- Sharing best practices for categorizing and identifying user agents (software and hardware used to download podcasts).
- Sharing data on bots to exclude from our measurement.
- Sharing best practices among sales teams for educating sponsors and agencies about podcast measurement in accordance with these guidelines.

See some of this work at the Podcast Measurement Wiki:

<http://podcast-measurement.wikia.com>

Contacts and Feedback

The working group encourages the open and voluntary adoption of these guidelines by all podcast producers, as well as by vendors, sponsors, agencies and other users of podcast data to facilitate common industry understanding of on-demand audio measurement. The members of the working group also welcome comments and feedback on these guidelines from other stakeholders, members of the public and other interested parties.

Stephen Haptonstahl, NPR - shaptonstahl@npr.org

Dan Jeselsohn, WNYC - djeselsohn@nypublicradio.org

Steve Mulder, NPR - smulder@npr.org

Eric Richardson, SCPR - erichardson@scpr.org

Kyle Wesloh, MPR/APM - kwesloh@mpr.org