SRG
STATION RESOURCE GROUP

The Audience Intelligence Project
Reporting on Our Work and Recommendations

Audience Intelligence is a multi-phase collaboration of leading public media organizations to gain insights about public media's digital audiences. Our long-term aim is to move from the aggregated audience estimates of broadcasting to individual, real-time data of the digital age, enabling public media organizations to serve individuals and communities more effectively and efficiently in their content choices, marketing messages, and fundraising.

Our goal for the recently-concluded first phase of the project was to tap existing data – at both individual stations and through a pooled dataset – to create first-cut digital audience profiles, while concurrently assessing and testing analytics tools, developing data sharing protocols, strengthening local analytics systems, and framing next steps for continuing growth in our capacity to generate digital audience insights.

As it turned out, this phase generated limited knowledge about our audiences – but rich insights about our organizations and the systems and tools we need to grow the use and value of our digital content.

We have shared many elements of our work and an emerging sense of the path forward with several hundred station and national leaders in conference workshops, retreats, and national planning meetings. We have seen a growing awareness of the immediate opportunities and long-term potential of advanced digital audience analytics for public media.

This report offers recommendations for next steps and reviews key findings from the work we have done so far.

Project partners for our first phase included Southern California Public Radio, Chicago Public Media, KQED, New York Public Radio, Oregon Public Broadcasting, St. Louis Public Radio and SRG. SRG provided overall project management. Lighthouse Datalab provided analytics and data science support. CPB funded this work.

Recommendations for Next Steps

The first phase of the Audience Intelligence Project produced important findings about the current state and significant challenges of public media's digital audience analytics.

We are also mindful of a larger context that surrounds this work, including portions of CPB's Digital Infrastructure Assessment, NPR's initiatives to develop and share digital capacities and similar work at PBS, and a metrics working group being coordinated through the Public Media Futures Forums in loose collaboration with NPR and SRG.

We see shared themes within these efforts:

- The need to continue building a culture of analytics and data-driven service at both the national and local levels of public media.
- The critical role of identity management solutions.
- A strategy of collaborative investments in high-cost infrastructure – technology, systems, software, and the personnel to develop and manage it – paired with distributed capacity-building at the local/regional level where audiences engage and contribute.
- A sense of urgency in the exploding territory of voice-controlled smart speakers and associated platforms and players.

We recommend a second phase of Audience Intelligence work that continues the discovery and piloting spirit of the first phase with proof-of-concept activity in select areas. This work should:

- Continue the model of a relatively small group of leading local public media organizations, working collaboratively, expanded to include NPR and PBS.
- Focus on working together to frame and facilitate best practices for management and sharing of digital audience data and advancing system readiness and support for large-scale investments and solutions.
- Demonstrate alternative identity management solutions such as a universal log-in/data handshake between stations and networks, and explore incentives and drivers for user registration.
- Continue development of stations' on-site digital audience analytics capacities, perhaps with case-study models of a fully-developed, station-focused system.
- Explore leveraging voice-controlled smart speakers and associated technologies as a new path to collect audience insights.

Key Findings

The accelerating pace of change in the digital media environment is an accepted reality. Yet public media organizations often perceive ourselves as working "against the backdrop" or "responding to the challenge" of such change, rather than embracing a dynamic in which we are deeply enmeshed and making choices that maximize our strengths in the marketplace.

How does the broader media environment shape our analytics work and larger public service goals?

- The use of digital analytics by our media competitors – already far ahead of public media – accelerated during the past year.
- Continuing rapid change in digital media platforms and devices, the latest being smart speakers and voice commands (Alexa, et al), underscore the need for swift tracking of evolving audience behavior and informed service responses.
- Enhanced capacities to generate meaningful audience analytics, along with actions based on the insights they yield, are fundamental to realizing public media's goals for relevant, engaging, widely-used, and sustainable services on digital platforms.
- Audience insights gained through digital analytics will also inform and shape our continuing broadcast services – by revealing users at the individual level, by tracking their interactions with our curated streams of content when they are offered on digital platforms, and by enabling compare-and-contrast analyses between broadcast-derived content and other digital assets.

This project has taken the partners closer to the front lines of work as fully-functional multi-platform media organizations and highlighted the distance we still need to travel. We have learned how we are increasingly entwined with technical giants – Amazon, Apple, Google, Microsoft, and others – which are both sources of tools, platforms, and devices on which we depend and competitors for our audiences' attention. We are working through a series of powerful inflection points. The significant disparities between our digital capacities and those of commercial media entities – of which audience analytics and business intelligence are no small part – represent a critical (some would say existential) threat to our public service goals.

Our work together has highlighted the importance of shared vision and collaboration in addressing large, complex tasks, as well as public media's limited skills and experience in successfully executing such solutions. We only scratched the surface in this territory, but enough to see both the need for leading organizations to seize the advantages of strong partnerships in the work ahead and the heavy lifting that will be required around business rules, governance, and system economics that will shift from broadcast-based to demand-based consumption patterns.

The data collection and analysis systems capable of fully meeting these aspirations are complex and sophisticated, requiring a higher level of technology and personnel investments than public media organizations have thus far been able to achieve.

Among the immediate significant challenges to public media's data analytics systems at both the local and national levels are:

- User identification/identity management. The vast majority of public media's digital users are anonymous, making it impossible to track users across the different platforms and devices with which they interact with a given organization or to track users across multiple organizations.
- Content tagging. Rigorous and standardized content tagging is an essential element of matching audience behavior with content preferences, both within organizations and across them.
- Data readiness. Public media audience databases are not yet ready for the advanced analyses we need to perform on a regular basis. Information remains siloed inside organizational departments, data integrity must be improved, data structures need to be better aligned to analytics goals, and hiring new staff with the skillsets needed to organize and mine the data remains a pressing challenge.

Here are some of the ways these challenges played out in the first phase of our work.

As we looked at the character of the available station data in spring 2017, almost all users were anonymous. This meant that we could not craft audience profiles along the familiar lines demography – who the users are – but would instead look to patterns of behavior – what the users do. We turned to another statistical technique – cluster analysis.

Cluster analysis classifies data points into groups based on *similarity*, in this case classifying digital users based on their behavioral attributes into groups of *similar* users. These groups are then studied to understand their characteristics and derive the representative profile.

But not all stations had data to support the technique. Lack of consistent tagging made it difficult to cluster users by content preferences. Conversely, some of the data fields that were

> Local analytics capacity is essential to inform local content, engagement, and membership.
>
> Centralized or federated approaches will help achieve efficiencies, scale, and standardization.

routinely available, such as the location of users, did not have enough variance to yield meaningful clusters within them.

Finally, stations tend to tailor their data analysis work to their particular interests, available data, and local analytics capacities. One organization may be throwing its data capacities into analyzing the podcast audience, another on membership, and another on website use.

These and other obstacles were frustrating, of course. But as we reminded ourselves on several occasions, identifying the problems helped us create a roadmap of the work to be done and our recommendations for next steps.

Overall, fully functional digital audience analytics for public media requires significant capacity building at both the national and local levels, buttressed by strong, structured partnerships between national and local organizations and among local organizations with significant digital activity.

- <u>Local/regional analytics capacity</u> is essential to inform the creation of local digital content and services and support the eventual real-time application of audience information to content offerings, local engagement, and membership development.
- <u>Centralized or federated approaches</u> to analytics infrastructure, technology, and data sharing will accelerate needed development and help achieve efficiencies, the power of scale, and standardization.

There is an apt parallel here to the development of public radio's journalism, which has grown dramatically through a combination of the power and leverage of national network service and the community connections, adaptability, and relevance of local newsrooms.

We have prepared a companion report, *Collaborative Audience Intelligence Solutions: Challenges and Opportunities,* which provides a more detailed accounting of the issues raised above and suggests possibilities for collaborative work going forward.

Digital Audience Patterns

Each partner station had the opportunity to "exercise" its own analytics systems, with support from Lighthouse Datalab, and to explore data about its digital audiences. While each station experienced its own insights and quirks, when we look across the partners we see some broad patterns that deserve comment.

Most striking is that our digital audiences appear to be quite transient. The percentage of digital users who showed up more than once within the month we examined was in the 15-25% range for the respective partners. The exception was New York Public Radio, which focused on

podcast downloads. While the percentage of users with more than one download – 50% – was higher than the web visitors examined by other partner organizations, this download rate is partially driven in part by podcast subscriptions and do not necessarily reflect actual listening.

The high percentage of one-time visitors overall is at least partially inflated by the lack of identity tracking – we cannot tell when a person returns on a different device. For example, a person who checks out headlines on their mobile phone one morning and streams a noon talk show later in the next day on their office desktop is recorded as two separate one-time users. We believe this is a minor rather than major factor.

We need to continue building a culture of analytics and data-driven service at both the national and local levels of public media.

Among users who do show up more than once a month, however, the number of sessions and page views jumps quickly, with averages among these repeat visitors suggesting weekly use and more than one page view per session.

All of the partners have significant audiences beyond their own communities – 20% to 30% of users are out-of-state for most partners. At New York Public Radio, for which the most successful podcasts are anchored to national programs like RadioLab, Freakonomics, and Snap Judgement, the percentage of out-of-state users is much larger. Several of the partners have significant broadcast coverage in more than one state; the out-of-state figures above reflect use beyond these areas.

Use of mobile devices is just edging out use on desktops at most partners. The exception again is NYPR, where podcast users prefer mobile devices over desktops by more than a 2 to 1 ratio.

Research Tools

We previously published a comprehensive assessment of digital audience research tools that we shared with SRG members and other system leaders. We know that some found this report too deep and detailed for their needs and expertise. But we urge public media decision-makers to engage with the infrastructure and technology of advanced audience analytics, at least at a broad outline level of functions and costs – in part to better understand what might make sense locally and what investments need to happen at some larger level. To that end we have prepared a *Tools Evaluation Summary,* presented as a companion document to this report, which we hope will provide an accessible entry point. The summary includes:

- Shared Audience Intelligence gaps and needs based on the assessments of all the partner stations

- A baseline set of metrics for each stage of the audience journey
- Proposed minimal sets of tools that fulfill project goals and needs, in two different configurations:
  - A system using largely "off-the-shelf" tools
  - A system based on a custom-built data warehouse
- A pro and con assessment (including pricing) of the two approaches
- A pro and con discussion of pursuing a shared data warehouse implementation approach

The Audience Intelligence Team

Southern California Public Radio
Bill Davis, President
Alex Schaffert, Director, Digital Media
Patrick Dougall, Audience Insights and Development Manager
Vijay Singh, Senior Product Manager, Mobile News & Audio
Bob Garcia, Project Manager
Tracy Irons-Georges, Institutional Giving & Grants Specialist

New York Public Radio
Laura Walker, President and CEO
Nate Landau, Senior VP and Chief Digital Officer
Peter Weingard, VP & Chief Marketing Officer
Kevin Stuart, VP, Digital Product & Technology
Valentina Powers, Senior Director of Digital Operations
Julia Schulhof, Senior Director, Digital Product Management
Eurry Kim, Manager, Data and Research

Chicago Public Media
Goli Sheikholeslami, President and CEO
Michel Ballard, Chief Digital Officer (until March 2017)
Kassie Stephenson, VP, Membership & Marketing
Beth Follenweider, Director, Business Intelligence
Thara Kollu, Senior Developer, Business Intelligence
Steven Rish, Manager, Audience Analytics

KQED
Tim Olson, VP, Digital Media & Education
Michael Lupetin, VP, Marketing & Brand
Andrew Alvarez, Senior Manager, Audience Insight & Ingenuity
Michele Murphy, Data Analytics Manager
Jessica Want, Senior Product Manager

Oregon Public Broadcasting
Steve Bass, President and CEO
Morgan Holm, Chief Content Officer
Jan Boyd, Director, Digital Strategy & Community Engagement

St. Louis Public Radio
Tim Eby, General Manager
Madalyn Painter Talla, Digital Media Manager
Sri Harsha Sankula, Audience Insights Intelligence Developer

Station Resource Group/SRG
Tom Thomas, co-CEO
Terry Clifford, co-CEO

Lighthouse Datalab
Sriram Subramanian, Principal
Ravi Ganta, Principal, Technology
Marish Gnanam, Analytics Project Manager
KP Pandy, Data Scientist
Alva Strand, Data Scientist