Coalition for Innovative Media Measurement

STRATEGIC REVIEW OF CONTENT MEASUREMENT

October 2024

Table of Contents

Introduction

The video ecosystem has experienced profound changes over the past decade as streaming consumption has grown. Today in the U.S., streaming constitutes over 50%[1] of total video consumption. Multi-Video Programming Delivery (MVPD) and virtual MVPD (vMVPD) pay TV subscriptions have declined from over 100 million U.S. households a decade ago to approximately 72 million in 2023.[2] The ARF DASH Study estimates that 62% of U.S. households subscribed to an MVPD or vMVPD service in 2023.[3] Today, over 95% of households have at least one streaming service subscription.[4]

The growth of streaming has caused equally profound changes in requirements for both ad currency and content measurement. While currency measurement that supports the ad ecosystem is transitioning to big data/ panel combined methodologies, content measurement has not kept pace with the migration of viewing to streaming and is still largely reliant on panels that were originally designed for linear TV.

Streaming content consumption datasets are often missing the needed program title- and episode-level detail due to the expense of coding them and transparency restrictions imposed by some streaming app owners. These content measurement gaps prevent programmers from answering their most pressing questions about content development, acquisition, distribution, discovery, monetization and other use cases.

The Coalition for Innovative Media Measurement (CIMM) commissioned this study to seek innovative solutions to fill the gaps, so that content measurement addresses their most pressing use cases, including measurement of viewer journeys across linear TV and streaming, using data that is comprehensive, title-level, persons-level and longitudinal data at scale.

1 The Gauge, The Nielsen Company, June 2024, https://www.nielsen.com/data-center/the-gauge/

2 S&P Global 2024 Trends in Global Media, November 2023, SPGMI_Preview_GlobalMedia_2024_FINAL.pdf

3 The ARF DASH Full Year 2023 Report, ARF-DASH-FULL-YEAR-2023_REPORT_FINAL.pdf (thearf.org)

4 Entertainment On Demand Q4 2023 US Barometer, Kantar, January 2024 https://www.kantar.com/north-america/company-news/us-streaming-services-must-focus-on-value-to-retain-subscribers-as-the-market-nears-saturation-point

About CIMM

The Coalition for Innovative Media Measurement is a non-partisan, pan-industry subsidiary of the Advertising Research Foundation, focused on cultivating and supporting improvements, best practices and innovations in measurement and currency, data collaboration and enablement, and the use of new metrics and approaches to understanding the value of media. CIMM embraces the entire media and advertising ecosystem and prioritizes effective collaboration to deliver meaningful change.

Every year, CIMM engages with its members to identify the most important and pressing priorities facing the industry. This study, a strategic review of the content measurement ecosystem for TV advertising, is an ongoing critical priority for our members. To find out more, contact info@cimm-us.org.

About the Authors

Gerard Broussard is Principal at Pre-Meditatedmedia (PMM) LLC, a measurement and impact advisory that helps advertisers, media agencies and publishers navigate the increasingly fragmented and digital world of marketing and advertising investing. PMM provides expertise in audience measurement, targeting, consumer segmentation and optimizing marketing spend to maximize ROI. Gerard can be contacted at gerard@pre-meditatedmedia.com.

Joan FitzGerald is CEO of Data ImpacX LLC, a measurement and marketing effectiveness consultancy that helps make the digital transformation happen in the media & entertainment industry. Data ImpacX provides measurement, optimization, marketing and software development services for publishers, agencies and brands. Data ImpacX helps publishers, agencies and brands succeed with data and analytics in an increasingly complex digital, TV, Smart TV, OTT, addressable and audio ecosystem. Joan can be reached at joan.fitzgerald@dataimpacx.com.

The views and opinions expressed in this report are solely those of the authors, and do not necessarily reflect the views or positions of any other individual or entity. This report is intended solely for informational purposes. Neither CIMM nor the authors make any representations as to the accuracy or completeness of any information contained in this report or in any report or web site linked to in this report. Neither CIMM nor the authors will be liable for any errors or omissions in this information or for any losses, injuries, or damages incurred from the display or use of this information.

A Note on Our Approach

This Report seeks to provide a strategic review of the content measurement landscape, identifying vendors, capabilities and gaps. In addition, this Report suggests recommendations for improvements to new and existing solutions, moving from solutions that can be delivered in a relatively short (6-12 month) time horizon to solutions that will require a longer time horizon to put in place. The starting point for this analysis was the CIMM Content Measurement and Analytics Working Group, which identified the most critical use cases for content measurement. To accelerate progress on content measurement, this Report was commissioned. The Project Steering Group provided extensive feedback and insights throughout the process. In addition, this Report is based over 30 in-depth hour-long interviews with executives from a wide range of companies supporting content measurement. The quotes provided throughout this Report are designed to be illustrative and derived from these interviews and meetings with the steering group. This Report should be read as representing the current state and best thinking as of September 2024.

The Project Steering Committee, including Daniel Butt, Lisa Heimann, Tom Hicks, Karen Miller, Don Robert, Melanie Schneider, Brian West, Chunguang Yu and Tom Ziangas, made critical contributions to this Report at each stage of its development, spending many hours and considerable effort sharing their expertise and discussing questions and concerns. Their insights guided the project from start to finish, and we are grateful for their careful consideration about content measurement. Without the leadership of Jon Watts at CIMM, this Report would not have been possible.

Executive Summary

In search of content measurement, programmers are urgently seeking solutions that advance the state of content measurement and enable more informed decisions about content development, acquisition, distribution, discovery, monetization and other use cases. Executives on CIMM's Content Measurement and Analytics project steering group expressed the need for data that would enable them to inform many use cases, including understanding the viewer journey – consumer viewing across linear and streaming titles and publishers over time and at scale. Equipping programmers with such data would help them to more effectively address their important content use cases.

Surveying the measurement vendor landscape, the project steering group recognized that no one vendor offered a comprehensive content measurement solution, and that multiple vendors would likely be required to fill content measurement gaps. To that end, in-depth interviews were conducted with over thirty vendors which provided feedback on their capabilities. Vendors included purveyors of data, audience measurement, content demand and valuation analytics, discovery services, metadata and others.

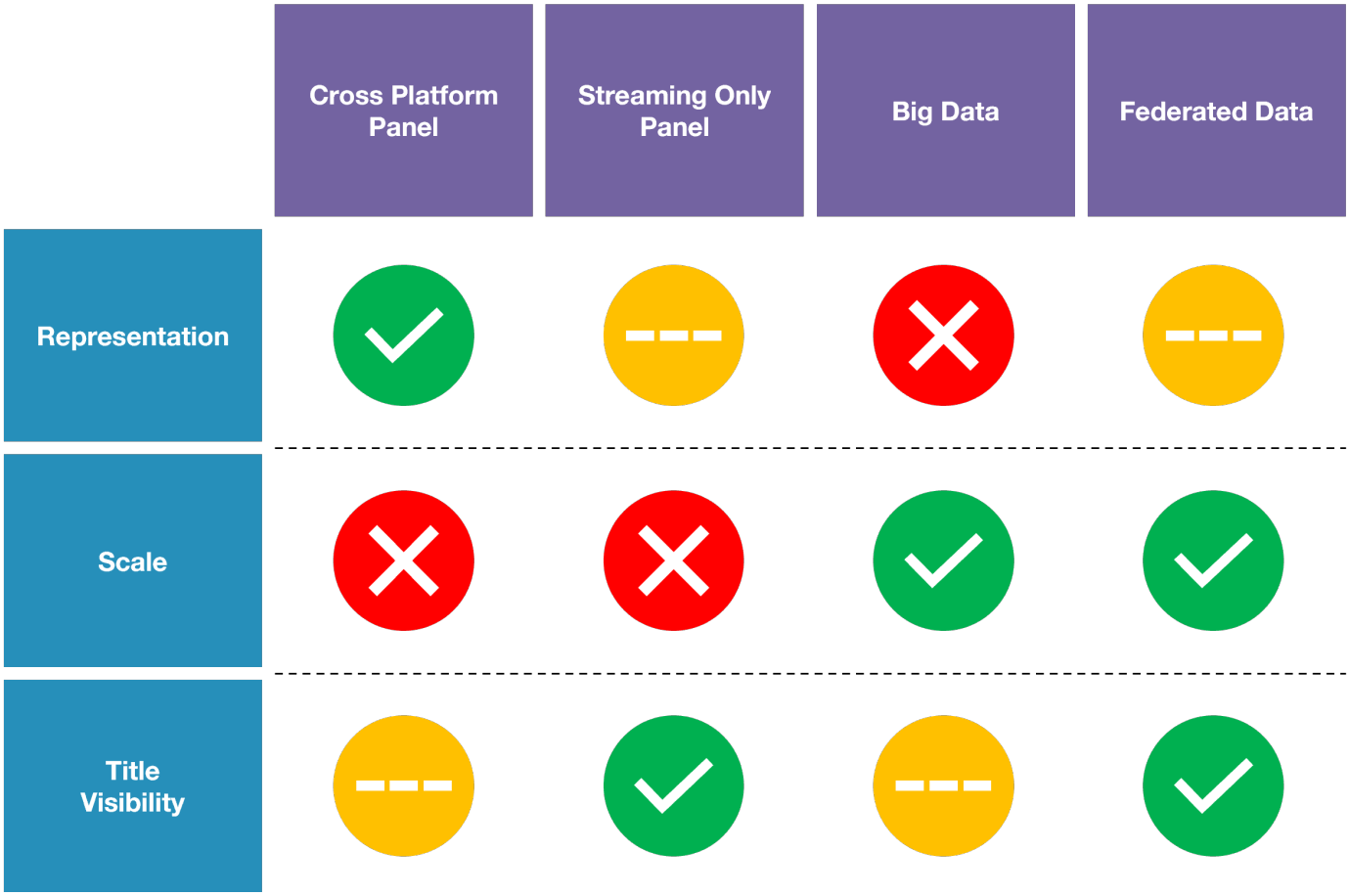
Executive Summary

Vendors identified barriers to content measurement**,** revealing three main challenges in capturing and reporting streaming content:

- **Economic -** incremental costs and resources required to expand panel sample size and increase the scope of ACR fingerprinting to measure streaming content, with FAST channels as an example.
- **Technical -** the desire for streaming app owners to permit additional ACR/fingerprinting, facilitate metadata identification and facilitate access to first- and second-party data.
- **Contractual -** many streaming app owners prohibit TV set manufacturers (OEMs) from capturing and reporting streaming consumption, which limits streaming content visiblity in Smart TV/ACR data.

Achieving representation, scale and content visibility were identified as important priorities to develop a solutions framework. Data sources that are representative across consumer groups (e.g., panels) don't necessarily possess scale, and solutions that have significant scale (e.g., big data) often lack representation. It is challenging to solve for both priorities using a single data source. Layered across the issue of scale vs. representation is the need to improve content visibility, achieved by attaching title, series and episode-level metadata to consumption data, especially in the streaming space. The chart shows the advantages and challenges of different data sources that could support content measurement. The advantages and challenges are complementary, so one data source alone won't solve for all content measurement requirements.

Data Building Blocks and Complementary Benefits



Executive Summary

The emergence of Streaming Only panels is an important development for filling in content-measurement gaps, because these panels can collect and report streaming publishers and titles thereby bypassing contractual reporting restrictions that exist between Smart TV/ACR data providers and streaming app owners. Enhancing Streaming Only panels by ensuring that they are representative and by increasing the number of streaming apps reported has the potential to improve visibility in measurement of streaming content consumption. Since these panels access viewing data from consumer-permissioned, subscription-based accounts, reporting of VOD consumption is the primary use case. The vendor's technology determines whether FAST channels (which don't require user accounts) and live viewing (that's not currently logged) can be captured.

> *We work with many streaming publishers, but for insights more than measurement. This report will help us put more rigor into our products so that they deliver even more value, especially to understand viewer journeys.*
>
> *– Streaming Only Panel Vendor*

Among big data sources, Smart TV/ACR meets more of the requirements for content measurement. Both Smart TV/ACR and Set Top Box (STB) data deliver scale, however, Smart TV/ACR data is poised to be able to meet more of the core requirements to measure both linear and streaming content in the future, if some of today's business and technical limitations can be solved. Smart TV/ACR can detect streaming signal while STB measures linear TV only. Smart TV/ACR also includes wider representation of homes by TV access where consumption habits can vary, most importantly broadband only (BBO) and some Over-the-Air (OTA) households. Smart TVs are in the majority of U.S. households compared to less than half of homes using an MVPD pay service.[5] Smart TV data providers are not prohibited from breaking out MVPD source, which is essential to understand the value of content to different distribution partners.

Despite these advantages, there are Smart TV/ACR limitations that need mention. Where Smart TV/ACR data does include streaming, it is limited to Over-the-Top (OTT) consumption from connected devices such as Apple TV, Chromecast, Fire Stick or Roku due to restrictions on data capture by streaming app owners. OTT consumption is estimated to be one-third or less of total streaming.[6] Smart TV/ACR data excludes native app streaming because ACR streaming data collection is prohibited by most app owners and because most streaming content is not presently included within ACR libraries.

For more information about how measurement vendors use Smart TV/ACR and Set Top Box (STB) data in their solutions, see CIMM's Convergent TV Measurement Guide[7] and CIMM's Smart(er) TV Data for Measurement report.[8]

This Report uncovered four crucial building blocks to address gaps in representation, scale and streaming visibility in an ecosystem where streaming and linear TV are both measurement imperatives.

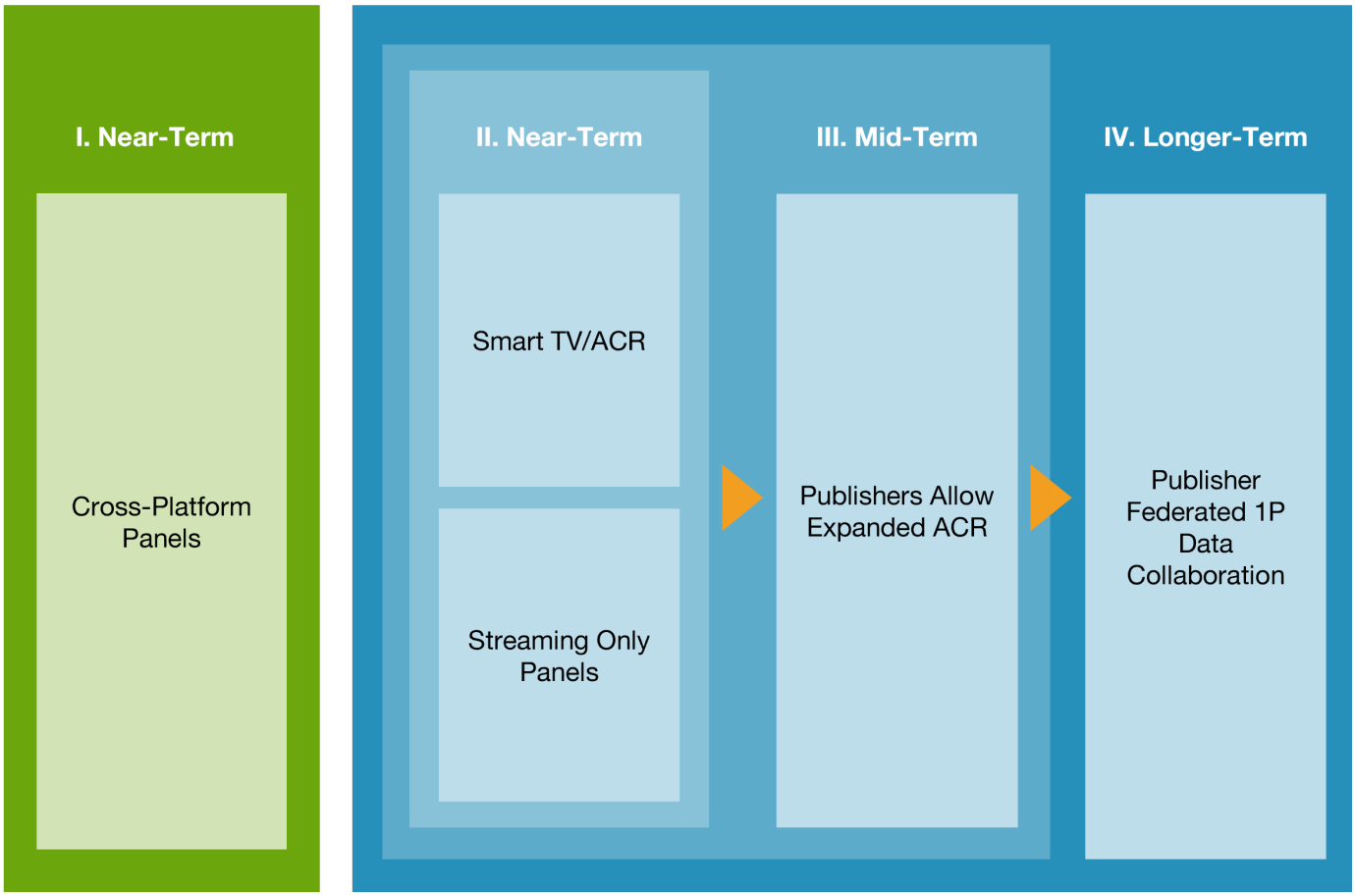
5 ScreenMedia US homes with traditional pay TV, vMPVD, and neither Q1 2024, https://nscreenmedia.com/us-pay-tv/

6 Estimated provided from proprietary research by Data ImpacX LLC https://www.dataimpacx.com

7 Coalition for Innovative Media Measurement CIMM Convergent TV Measurement Guide 2023, cimm-us.org

8 Coalition for Innovative Media Measurement CIMM Smart(er) TV Data for Measurement Report, February 2024, CIMM-Smarter-TV-Data-2024.pdf (cimm-us.org)

Phased Solutions to Improve Content Measurement



I. Near Term, Cross-Platform Panels

The existing Nielsen cross-platform panel used for currency measurement continues to be an important building block for content measurement, delivering representation, including consumer groups not represented in big data. For example, cross-platform panels include consumption from broadband only (BBO) households using TV sets not connected to the Internet and Over-the-Air (OTA) only households that are not in Smart TV/ACR data. The ARF DASH Study estimates that 34% of U.S. households are BBO and 14% have OTA access to TV/video.[9] However, like all panels, achieving greater scale and capturing long-tail consumption is a continuing challenge. A detailed analysis of panel strengths and limitations can be found in a recent CIMM report "The Future Role and Value of Panels in U.S. TV Measurement".[10]

II. Near Term, Smart TV/ACR and Streaming Only Panels

Near-term, third-party big data supplemented with panel data can solve for scale, representation and content visibility in content measurement. Smart TV/ACR data has sufficient scale to detect "long-tail" viewing, especially for linear TV, that falls prey to the sample size limitations of a panel. Smart TV/ACR data sources that measure both Over-the-Top (OTT) streaming and linear TV provide important signal for understanding cross-platform content consumption.

> *The report identifies streaming data sources and services we frankly didn't know about. If we did know about them, it's clear we can improve them to use them across more of our use cases.*
>
> *– Major Publisher*

Streaming Only panels can fill in the streaming gaps in big data, measuring virtually all VOD streaming publishers and titles, unencumbered by restrictions that prohibit reporting in Smart TV/ACR data. Streaming Only panels require improvement to be representative, and Smart TV/ACR data requires improvement to capture more streaming, but both in combination hold the promise of solving for near-term content measurement challenges.

9 The ARF DASH Full Year 2023 Report, Advertising Research Foundation, 2023 ARF-DASH-FULL-YEAR-2023_REPORT_FINAL.pdf (thearf.org)

10 The Future Role and Value of Panels in U.S. TV Measurement, Coalition for Innovative Media Measurement (CIMM), Oct 2024, cimm-us.org

III. Mid-term, Publishers Allow Expanded ACR

Coverage in Smart TV/ACR datasets could be more comprehensive if streaming app owners granted permission for more ACR data collection of streaming content by data and measurement providers. In this approach, streaming app owners would allow Smart TV/ACR data providers to ACR/fingerprint, collect and report title-level streaming consumption data across both OTT and native app streaming. To be sure, granting permission for more streaming coverage also requires business agreements. However, there is expected to be less complexity because, unlike a full data collaboration framework, data ownership and data collection is the responsibility of the third-party data provider, not the streaming app owner.

IV. Longer term, Publishers Engage in 1P Data Collaboration

Publishers can achieve even greater accuracy, scale and content visibility through federated data collaboration. “Federated” data collaboration is where publishers provide access to first-party streaming data from their owned streaming apps to achieve greater streaming measurement accuracy and title-level visibility. Such a collective undertaking would require a system of governance, access rules, privacy, security and technology to enable measurement sourced from first-party streaming data.

Note that publishers do not necessarily own all of the streaming apps through which their content is distributed. When they do own the app, the publisher also owns the first-party data for the streaming content. If the streaming content is distributed through an app owned by a distribution partner or content licensee, the distribution partner or content licensee owns the first-party data. Depending upon the relationship between the distribution partner or licensee and the publisher, the publisher may have access to the distribution partner or licensee’s first-party data (which is the publisher’s second-party data). If the publisher has negotiated rights to the second-party data, it is typically aggregated data. If it is user-level, the user data is anonymized.

Executive Summary

Federated data collaboration should include both consumption data and metadata. This Report identifies the promise of data collaboration and provides more detail on streaming *consumption* data collaboration. However, ideally, the collaboration will include both streaming consumption data and *metadata*. Lack of naming standards in metadata, including distributor, episode, series, title names and identifiers, and prohibitions on passing and reading metadata are significant challenges in the data supply chain, causing individual publishers, data providers and measurement providers to establish their own watermarks, fingerprints, dictionaries, algorithms and processes to identify what the content is and whether the content is the same across different distributors. Adoption of transparent, industry-wide metadata collaboration standards could help ensure accurate, consistent content identification. A CIMM study exploring the benefits and challenges of industry-wide watermarking standards is currently underway and expected to be published in late 2024.

Programmer decisions about content are complex and cover a wide range of use cases including creation, licensing, distribution, windowing, valuation and other critical decisions. Publishers realize no one solution can deliver all content use cases, and that achieving their vision of measuring viewer journeys – based on data that are representative, at scale, persons-level, longitudinal and title-level streaming and linear TV - sets a very high bar for delivery.

This Report identifies key building blocks required to establish a more robust set of content measurement solutions. We believe that solutions based on third-party building blocks can be improved and made available in a relatively short time horizon. Solutions that require federated data collaborations across publishers will deliver more benefits but are more complex and expected to take more time. We urge our programmer partners to prioritize partnerships with content measurement vendors and to support and participate in a federated data collaboration framework to make more robust content measurement a reality for themselves and the industry.

Content Measurement Defined

Major U.S. publishers convened a Content Measurement Working Group through the Coalition of Innovative Media Measurement (CIMM). They collaborated to define content measurement, and to identify the main use cases that would help measurement providers understand their needs.

The Working Group defined content measurement as helping programmers *understand from the point of view of the viewer who, how and when audiences are consuming video content, by episode and title across all distribution platforms, over time.* They distinguished Content Measurement from Ad Measurement, defining ad measurement as focused on quantifying exposures to advertising impressions and campaigns, helping advertisers to understand business outcomes.

CIMM CONTENT WORKING GROUP

What do we mean by Content Measurement? And how is it different from Ad Measurement?

Content Measurement

Helps **programmers** understand from the **POV of the viewer** who, how and when audiences are **consuming video** content, by **episode and title across all distribution platforms, over time.**

Ad Measurement

Focused on quantifying exposures to advertising impressions and campaigns, helping advertisers to understand business outcomes.

Use cases identified by the working group for content span a wide spectrum, including the following that are detailed in the working group report:

- Content development, scheduling and windowing strategy,
- Business development and monetization strategy,
- Direct-to-Consumer (DTC) customer acquisition, retention and marketing, and
- Supporting external business partners and communications

> *The Working Group's description of Content Use Cases is incredibly valuable. It's become our roadmap to guide our product plans and improve content measurement*
>
> *– Measurement Vendor*

Industry Effort for Content Measurement

What use cases are we solving for?

Content Development, Scheduling & Windowing Strategy

- What content do we need?
- What content is working and what is the "White space"?
- Produce or acquire it?
- How can we maximize the lifetime value of our studio content?
- How do we support DEI initiatives?
- When should we pick up an additional season vs end series?
- Where do we put our content to grow and maximize audiences?
- When should it debut, how should it roll out, and what's the impact exclusivity vs broader exposure?
- What is the customer journey for content discovery and later viewing?
- How do viewers move across platforms?

Business Development & Monetization Strategy

- What are our marketplace opportunities, and how do we understand our position relative to competitors?
- Where should we look for future licensing, M&A, and other deals?
- How do we maximize the value of our networks, stations, apps, and content to our distribution partners (both linear and digital)?
- How can we optimize ad sales revenue thru the planning and forecasting processes?

Supporting External Business Partners & Communications

- How do we measure, manage and position our viewership for talent, profit participants, and other rights holders and contracts?
- What metrics are useful for benchmarks in which deals?
- What information helps highlight the strategies we're pursuing, our successes, and the value the company is delivering for our investors?
- How can we best manage our public narrative with the press?

DTC Consumer Acquisition, Retention & Marketing

- How can we acquire and retain subscribers on our streaming service, while building reach on both ad-supported and ad-free tiers?
- How do we make viewers aware of our programming and make sure they show up to watch?

Content Measurement Requirements

CIMM's Content Measurement and Analytics Project Steering Group is comprised of leading publishers who sought a strategic review of content measurement vendors, capabilities and gaps. This effort resulted in two major deliverables:

- A **"CIMMscape"** which graphically displays the contributing components of the content measurement marketplace in the U.S. and vendors who support each component.
- A **Strategic Review** of content measurement building blocks, capabilities and gaps, and a series of recommendations to address the gaps to foster more useful, relevant content measurement services.

Requirements that Prioritize Viewer Journeys

1. Meet main use case requirements
2. Feasible and pragmatic
3. Achievable timing
4. Phased approach
5. Affordable cost

→

- Persons
- Streaming and linear TV
- Title, series and episode-level visibility
- All devices and access methods
- Syndicates reporting
- Over time
- Representative
- At scale

There are many use cases for content measurement, and different use cases can be supported by different data. To help focus this effort, the Project Steering Group prioritized data to understand consumer viewing journeys. Viewer journeys was prioritized because it represents the most complex content analysis with significant requirements. Plus, the data for viewer journeys analysis can be applicable to other less complex use cases to help inform content production, acquisition, distribution, scheduling and windowing decisions, among others.

> *We need more than data about popularity of streaming shows. Content measurement has to be consumer-focused, collecting consumption from the same people over time. We focused on viewer journeys because it's the most challenging use case.*
>
> *– Major publisher*

The data required to achieve viewer journeys analysis should be persons/consumer-level, include streaming and linear TV, has visibility at the title, series and episode-level, covers all devices and access methods, is reported on a syndicated basis, measures consumption over time/longitudinally, is representative and has scale.

In other words, longitudinal data measuring individual users across the entire video ecosystem.

This is indeed a tall order, but to understand how consumers make their choices, programming preferences across the competitive landscape, audience flow, tune-in/tune-out and performance of titles by platform, device and audience among other questions about content consumption, requires longitudinal, user-level data.

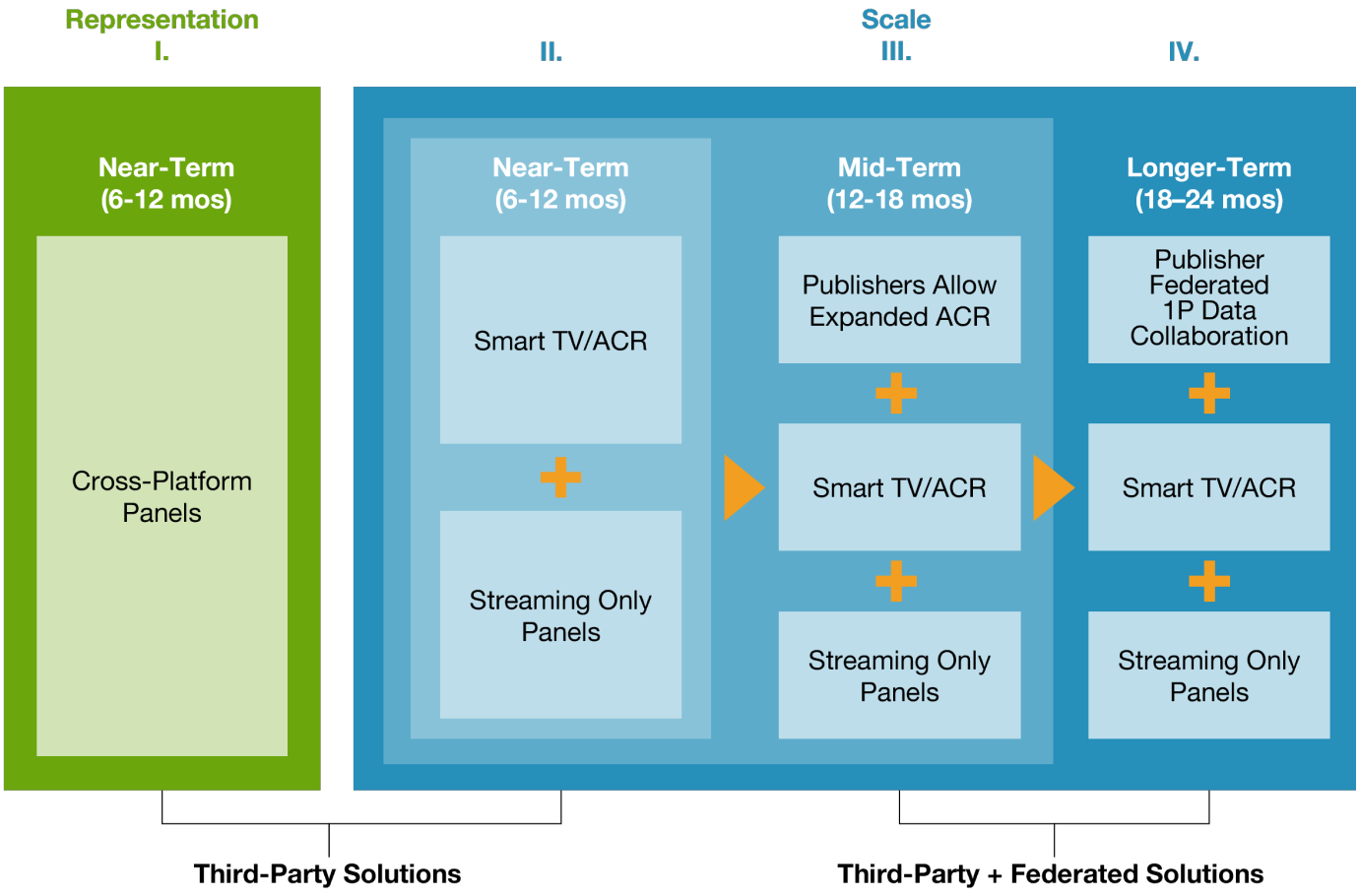
There are other content measurement use cases that do not require user-level, longitudinal data. These use cases can be supported by other types of data including aggregated, non-user level data, self-reported survey data, and data collected for other non-measurement purposes. They are described in the Other Use Cases section of this Report.

Content Measurement Solutions Architecture

Recommendations for a content measurement solutions architecture include four building blocks to address the priorities of representation, scale and content visibility. They are designed to provide a structure for creating increasing value in content measurement as they are activated and improved over time. Increases in scale, accuracy and content visibility are achieved in each ascending step from building block I to IV.

The first two options are based solely on third-party solutions, which have a faster speed-to-market than federated solutions because streaming app owner cooperation is not required for data collection and reporting. The third and fourth option combine third-party solutions with federated approaches and are expected to have a longer time horizon.

Content Measurement Solutions Architecture



I. Continue to improve cross-platform panels, including the incumbent current panel as well as new panel technologies emerging in the marketplace.

Cross-platform panels are representative and have other advantages, and can continue to be improved over time, for example, to collect streaming consumption from more panelists, to capture title-level visibility from mobile and PC devices and to distinguish between ad-supported and non-ad supported content. There are new AI-based technologies that have emerged that could reduce costs and improve content identification in cross-platform panels.

II. Increase streaming coverage in Smart TV/ACR Data and supplement with Streaming Only panel data.

Smart TV/ACR data can achieve significant scale; however, coverage of streaming apps and titles needs improvement. Some Smart TV/ACR data providers don't measure streaming at all. Others measure only OTT streaming, but the measurement is not currently comprehensive of OTT consumption. Comprehensive streaming – even if it is only OTT – would be a significant improvement. In addition, enhancements such as weighting by OEM, accounting for co-viewing and identifying MVPD are important to maximizing utility of Smart TV/ACR data for content decisions.

To be sure, Smart TV/ACR data will continue to have gaps, especially the absence of native app streaming and mobile/desktop streaming consumption among others. Streaming Only panels are a data resource that, with improvements such as greater representation and inclusion of more streaming apps in measurement, have the potential to fill in the streaming publisher and title gaps in Smart TV/ACR data. Streaming Only panels are a unique, lower-cost approach to measuring streaming consumption across virtually all publishers and devices, including Smart TVs, mobile phones and PC/desktops.

In Streaming Only panels, panelists grant research/measurement firms permission to access their streaming account data, either by downloading streaming logs or using an app from the research/measurement provider to access the data. Although an economic analysis is beyond the scope of this Report, Streaming Only panels appear to be lowercost and have less panelist burden than crossplatform panels. Most cross-platform panels used for currency measurement globally require meter installation on multiple sets and ongoing interaction to indicate who is viewing. In Streaming Only panels, once the participant grants access to their streaming logs, there is less required activity by the panelist. However, Streaming Only panels don't collect linear TV data and today they are not necessarily representative. Whether Streaming Only panels collect FAST and live streaming depends upon the technology used by the provider.

III. Smart TV data coverage could be further expanded if streaming app owners grant permission for more ACR/fingerprinting.

This step would improve streaming coverage in existing Smart TV/ACR data and could be a first step towards a federated approach across publishers. In this scenario, streaming app owners allow third-party Smart TV data providers to ACR/fingerprint their streaming content, helping to fill in streaming gaps, especially for native app streaming that isn't measured in Smart TV/ACR data today. The advantage is scale but with potentially less complex business agreements compared to a full data collaboration framework, because the third-party data provider, not the streaming app owner, is responsible for data collection.

IV. Streaming content measurement will be more accurate with greater scale when publishers establish a Federated Data Collaboration Framework.

A Data Collaboration Framework would enable measurement sourced from streaming app owner first-party data. We envision a Data Collaboration Framework will require a longer time horizon to put in place than third-party solutions because the governance, access rules, policies and business agreements that are needed are quite complex. Regardless, this level of cooperation and participation by streaming app owners is necessary to ultimately achieve more accuracy, scale and greater control among publishers for content measurement.

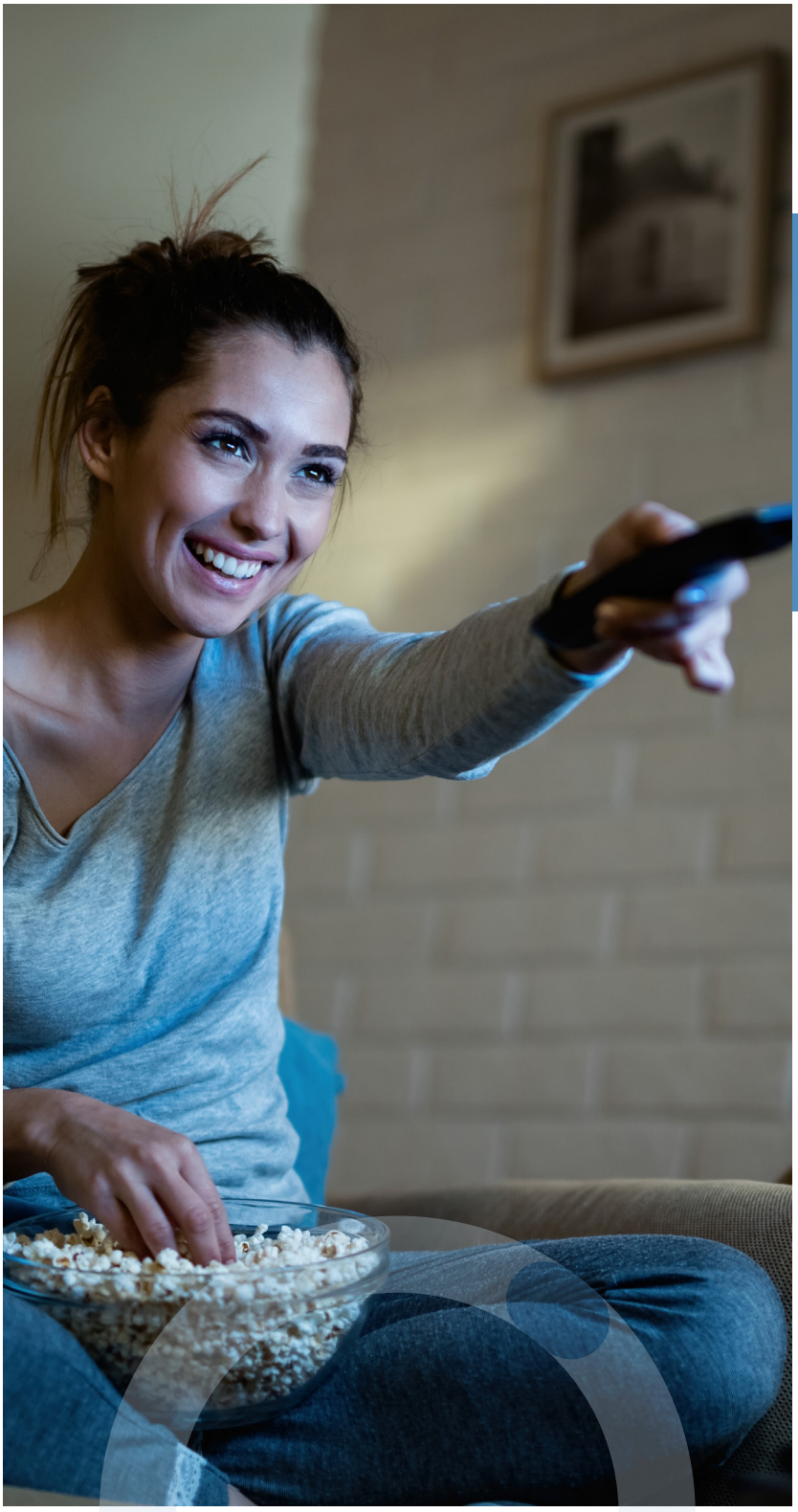
A comparison across data sources and their applicability to content measurement requirements is in the Appendix.

Even when federated approaches are adopted and the advantages are realized, we envision third-party solutions, including big data, Streaming Only panels and Cross-Platform panels, as critical components of content measurement.

Panels are critical because they can measure streaming regardless of cooperation or participation by streaming app owners, providing essential competitive visibility. Cross-platform panels provide a combined linear TV and streaming perspective but, unlike Streaming Only panels, they don't typically report mobile/PC consumption at the title-level. We view both Cross-Platform and Streaming Only panels as essential resources for content measurement.

Cross-Platform and Streaming Only panels have limitations. As panels, they are subject to statistical error that occurs in all panels which means measurements will not perfectly match census-based metrics that could be available from federated first- and second-party data. These limitations may spur participation by streaming app owners in federated solutions to improve accuracy and scale compared to third-party solutions alone.

The Future of Federated Data Collaboration



A federated first-party data collaboration framework has substantial benefits for participating streaming app owners, and we are optimistic that issues and concerns can be resolved to make a federated data collaboration a reality.

> *If we can move forward on data collaboration, that would be an outstanding outcome of this project. We'd love to see a framework for publishers to allow more data collection and to share metadata with us. That would really improve measurement.*
>
> *– TV/Streaming Data Licensor*

- First-party streaming consumption data is census-level, and streaming app owners can achieve greater scale and accuracy than is possible from panel or Smart TV/ACR data alone.
- Federated data at the title-level would improve content visibility across participating streaming app owners compared to Smart TV/ACR data where native app streaming reporting is restricted.
- The framework enables streaming app owners to exercise greater control over measurement and decision-making, which can be beneficial for streaming app owners but also challenging to manage across publishers with different needs and viewpoints and different levels of first-party streaming data ownership and access to second-party streaming data.

The Future of Federated Data Collaboration

Federated approaches to collaborate on first-party data face barriers, but they are business-related, not technical, and not insurmountable. The good news is that there are few technical barriers to federated data collaboration. There are many technologies to ensure data security and privacy including data clean rooms and identity solutions. There are best practices that have been used for decades to ensure compliance with Video Privacy Protection Act (VPPA) and other privacy regulations. ACR/fingerprinting technology is mature and well-understood. There is a U.S. Joint Industry Committee (JIC) working to enable collaboration on ad campaign performance data, and the technical underpinnings could be used by a content measurement data collaboration.

The barriers to federated data collaboration are business-related, requiring new business agreements, relationships and operating standards to succeed. We envision a Data Collaboration Framework where streaming app owners agree on crucial issues to make measurement sourced from their first-party streaming app data a reality, including governance, access, reporting standards, among other negotiations.

Data collaboration is such an important topic to future-proof content measurement. This report will help us publishers move forward, especially to work through the very detail questions that are needed to make it work.

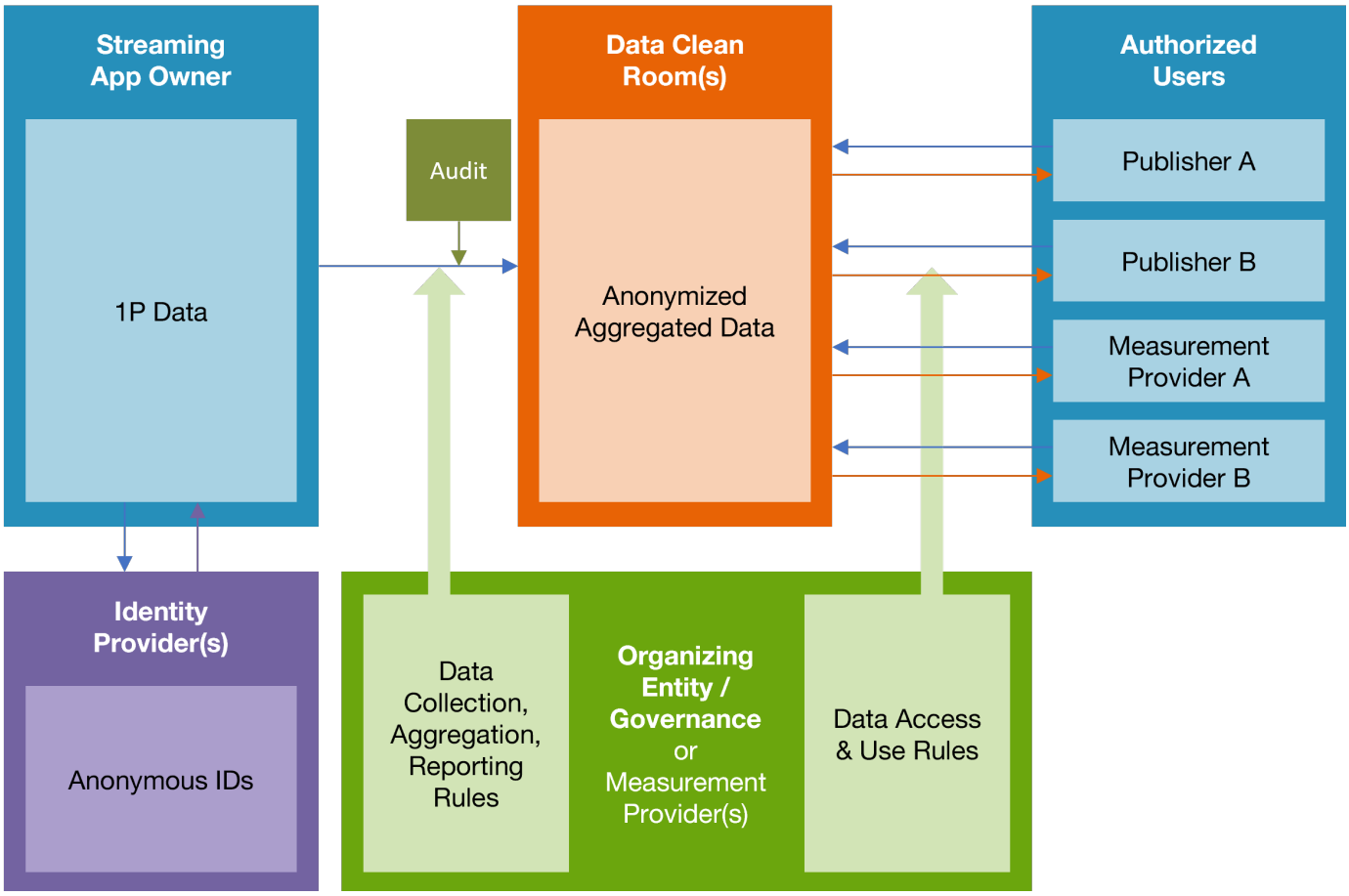
– Major Publisher

A federated approach for content measurement has additional requirements beyond the JIC to establish rules and definitions that guide content data collection, aggregation and reporting, making content data collaboration more complex than the ad campaign performance collaboration that is currently being undertaken by the JIC:

- Compared with ad impressions data (which are well-understood and already frequently shared between publishers and ad buyers), content data will require data collection, aggregation and reporting standards and a higher level of curation to distill the data into a form that publishers are willing to share, such as title, date/time and duration of consumption.
- The commercial arrangements between ad buyers and publishers incentivize publishers to share their first-party ad impressions data. Content measurement has no such commercial arrangement, and the risk/benefit analysis is more complex when potentially sharing data with a competitor.

An example of a federated data collaboration framework is outlined below, where the main constituents include publishers, identity providers, an organizing entity, auditors and technology/measurement providers.

Model for Federated 1P Data Collaboration



- **Publishers (streaming app owners**) own their first- party consumption data from their streaming apps and frequently have rights to second-party data from their distribution partners or streaming content licensees. However, second-party data is often aggregated, not user-level. Publishers that distribute content through their owned streaming apps will have first-party data. Publishers that license and distribute content through other parties/apps could have second-party data, depending upon their negotiations with the other parties.
- **Identity providers establish anonymous IDs** and can be a source of demographics and consumer attributes. Publishers send their respective user account/PII data (without consumption data) to one or more identity provider(s). The identity provider(s) match the account/PII data to their identity spine and assign a persistent anonymous ID which is sent back to each publisher for the publisher's users/accounts.
- **An organizing entity (which could be one or more measurement providers)** establishes the governance and process by which publishers collaborate. Under the auspices of the organizing entity (or measurement provider(s)), publishers establish the rules and standards that govern the data collaboration, including the six dimensions in the table below.

Governance Functions of the Data Collaboration Organizing Entity

Function	Definition	Objective
Data Workflow	The map of the expected, ongoing data and workflow between publishers, identify providers and other vendors in the collaboration.	Used by the organizing entity to ensure ongoing, timely workflow between publishers, identify providers and other vendors.
Data Aggregation	Data processing rules, including defining duration, defining the viewing time required for the consumer to be considered a viewer, defining data aggregation rules.	Used by streaming app owners (or their technology vendors or measurement providers) to ensure datasets are created using the same methodologies.
Data Reporting	Establishing requirements around reporting granularity and frequency.	Used by streaming app owners (or their technology vendors or measurement providers) to ensure standardizes in reporting cadence.
Data Access	Identifies the publishers and/or their vendors which are allowed to access the data, the time period that access is granted, and the datasets or data within each dataset that each is allowed to access. For example, a publisher may have access only to data for their viewers and not data for the entire universe of viewers.	Used by the organizing entity to manage and control all aspects of data access by publishers and/or their vendors.
Data Use	Defines the permitted use cases for each dataset (i.e, how the data can be used, such as internal use only and not for ad sales). This could include review/approval of models and algorithms used against the data.	Used by the organizing entity to manage and control how the datasets are used by publishers and/or their vendors.
Data Privacy and Security	The organizing entity will require that members adhere to data privacy regulations and data security standards.	Ensures that participants in the collaboration adhere to data privacy regulations and provides participants a framework to resolve data security concerns.

- **Auditors and audit processes** ensure that data provided by participating publishers follows the agreed-upon data processing and aggregation rules.
- **Privacy and security technologies** are the wide range of technologies such as data clean rooms provided by a wide range of vendors.

CIMMscape: The Content Measurement Vendor Landscape

Modeled after a "LUMAscape,"[11] the Content Measurement CIMMscape is designed to organize companies in the content measurement industry into specific industry categories in a single view. The CIMMscape was constructed in recognition of the fact that no one firm can currently deliver a comprehensive content measurement solution. An understanding of the entire content ecosystem is critical to identify key categories and firms that could advance the state of content measurement.

> *The CIMMscape gives us a snapshot of the vendor landscape and the component parts that need to go into content measurement. It's a great starting point to understand the overall industry and who the players are.*
>
> *– Major Publisher*

The CIMMscape should be considered a work-in-progress and may not include all companies within each category. We welcome your feedback to improve it over time.

11 A Visual Guide to the Digital World, LUMA Partners, https://lumapartners.com/lumascapes/

CIMMscape: The Content Measurement Vendor Landscape

PRODUCTION

Production

20th Television
ABC Signature
A&E Studios
Amazon Studios
AMC Studios
CBS Studios
Crown Media Productions
Fox Entertainment
Lionsgate
Paramount
Sony Pictures
Universal
Walt Disney Productions
Warner Bros

Distribution

OTA

ABC
CBS
CW
FOX
ION
NBC
Estrella TV
Telemundo
Univision
MyNetworkTV
PBS

Streaming OTT

Amazon Fire
Apple TV
Google Chromecast
Roku

Streaming OEM

Hisense
LG
Phillips
Samsung
Sanyo
Sharp
Sony
TCL
Toshiba
VIZIO

MVPD

Charter/Spectrum
Comcast/Xfinity
Cox
Dish
DirecTV
Hughesnet
Mediacom
Optimum
RCN
Verizon FIOS
WOW

Streaming App

A&E
Amazon Prime
AMC Plus
Crackle
Disney+
Hallmark
Netflix
Paramount+
Peacock
YouTube
Vevo

Streaming FAST

Cineverse
Freevee
IMDb
LG Channels
Pluto
Roku Channels
TiVo+
Tubi
Samsung TV Plus
Xumo
VIZIO WatchFree+

Theatrical

AMC
Regal
Cinemark
Cineplex
Marcus

Discovery

EPG/UX

Comcast
TiVO
ThinkAnalytics

ACR

CCRMedia
Utelly

Apps

Whip Media
Younify

Streaming/Measurement Tech

Adobe
Amagi
Conviva
Kochava
RealityMine
Dalet
Frequency
Wurl

Authentication

Adobe
C2PA
Microsoft

Data Licensors

Panel

Linear TV DVR, VOD, OTA & Streaming

HyphaMetrics
Nielsen
TVision

Streaming VOD

Digital i
Measure Protocol
Plum Research
YouGov

Big Data

Linear TV DVR VOD

Charter/Spectrum
DirecTV
Comcast/Xfinity
Dish
FourthWall Media
LG Ads
Samba TV
TiVo
VIZIO/Inscape

Linear TV OTA

LG Ads
Samba TV
VIZIO/Inscape

Streaming OTT

Samba TV
VIZIO/Inscape

Theatrical

Comscore
IMDb
TMDB

Identity Demos & Audiences

Audience Acuity
Acxiom
Epsilon
Experian
The Trade Desk
TransUnion

Metadata & Schedules / Windows

Ampere
Fandango
Gracenote
Hive
iMDB
iSpot TV
Kinetiq
Luminate
Parrot Analytics
Rotten Tomato
Stream Metrics
TiVO
TMDB
ThinkAnalytics
Who's Watching TV

Sample & Surveys

DISQO
Dynata
Kantar

TV / Streaming Measurement

Comscore
HyphaMetrics
iSpot TV
Kantar
Nielsen
Samba TV
Tvision
VideoAmp

Streaming Viewership & Viewer Journeys

Digital i
Measure Protocol
PlumResearch
YouGov

Streaming Demand Valuation & Insights

Cinelytic
Luminate
MarketCast
Muso
Management Science Associates
Parrot Analytics
ScreenEngine/ASI
Stream Metrics
Younify
Whip Media

MONETIZATION

CIMMscape: The Content Measurement Vendor Landscape

The CIMMscape includes the following elements:

- **Production** - Leading studios and production companies that create video content.
- **Distribution** - Content distribution points as experienced by consumers, including OTA, MVPD, Apps etc.
- **Streaming Tech/Authentication** - This section contains examples of companies that enable streaming distribution, data collection and authentication.
- **Discovery** - Companies that provide EPG (electronic program guide), UX (user experience) and consumer apps for consumer discovery of content.
- **Data Licensors** - Companies that license syndicated, anonymized, user-level data about TV/video consumption, including data collected from consumer research panels and big data sources such as STB and Smart TV/ACR data.
- **Identity/Demographics/Audiences** - Companies that enable user-level matches between TV/video consumption data and consumer attributes, interests and other data sourced from their ID graphs.
- **Sample & Surveys** - Companies that provide sample for panel and surveys by other companies. These companies also collect data via surveys.
- **Metadata & Schedules/Windows** - Companies that provide descriptive content data such as title, episode, genre, production year, talent as well as the timing of content availability on platforms and apps.
- **TV/Streaming Measurement** - Companies that have established methodologies to curate linear TV or linear TV and streaming consumption data to report syndicated estimates of audience size and composition and analyze viewer journeys for ad campaigns and/or content.
- **Streaming Viewership & Viewer Journeys** - Companies that collect streaming consumption data and have established methodologies to curate it to report syndicated estimates of audience size and composition and analyze viewer journeys for streaming titles.
- **Streaming Demand, Valuation & Insights** - Companies that provide insights about streaming consumption and measure demand and forecast viewing potential for specific programs and program genres to inform content strategy.

Vendor Perspectives

In-depth interviews were conducted with a wide range of vendors in the content measurement landscape. Interviews focused on current capabilities and whether they are able to capture viewer journeys on both linear TV and streaming.

We appreciate the participation and perspectives of all of companies we interviewed, including C2PA, Cinelytic Inc., Comscore Inc., Digital i Ltd., Dynata LLC., FOX Corp., Gracenote Inc., HyphaMetrics a HyphaNetwork Company, Inscape a division of VIZIO Holding, iSpot.tv Inc.; Kantar Group Ltd., LG Ads Solutions a division of LG Electronics Inc., Luminate Data LLC., Management Science Associates Inc., MarketCast LLC., Measure Protocol Ltd., Muso TNT Ltd., Nielsen Media Research Inc., Open AP, Parrot Analytics Ltd., Plum Research Sp. z.o.o., Reality Mine Ltd., Samba.TV Inc., Screen Engine/ASI LLC., Stream Metrics Inc., Think Analytics, Ltd., TiVo a division of Xperi, TVision Insights Inc., VideoAmp Inc., Whip Media Group Inc., YouGov PLC. and Younify Inc.

Vendor Perspectives

Obtaining vendors' perspectives on the challenges in improving content measurement provided important insights to informing solution alternatives for content measurement. Vendors cited three main barriers to improving content measurement:

1. **Economic** - Vendors mentioned incremental cost to invest in panel expansion and/or representation. For example, some consumer segments may require separate measurement such as streaming only and OTA. For both panel and big data approaches, the cost and labor intensity of expanding ACR fingerprinting to measure more long-tail streaming, including FAST channels, represents a challenging undertaking with vendors believing there is limited ROI attached to the investment required.
2. **Technical** - Vendors expressed the desire for streaming app owners to cooperate with and enable data collaboration for syndicated measurement. Some suggestions for streaming app owner actions include enabling metadata identification, permitting fingerprinting, allowing pixeling and enabling access to first- and second-party streaming consumption data.
3. **Contractual** - Vendors cited restrictions in their agreements with streaming app owners, such that they are not permitted to report content at the title-level. These agreements play a significant role in content measurement opacity. For example, capturing and reporting viewing data consumed directly via the publisher's streaming app (i.e., "native app" consumption) is prohibited within many agreements between OEMs and streaming app owners. MVPDs prohibit viewing breakouts of their subscriber footprints. Meanwhile, some vendors aren't licensed to use data for viewer journeys or content use cases.

Content Use Cases Beyond Viewer Journeys

This Report focuses on use cases that require user-level, longitudinal data to understand viewer journeys, and uncovers a relatively new data building block referred to as “Streaming Only panels” where streaming consumption data is captured from panelist streaming account logs. Panel sizes range from thousands to tens of thousands, and the data is typically collected on a global basis. These data enable publishers to make comparisons across different countries to see what is trending and when. Streaming Only panel providers include Digital I, Measure Protocol, Plum Research, YouGov and Younify.

There are many use cases – identifying ‘white space’, forecasting demand, valuing content and others – that do not necessarily require longitudinal, user-level data about viewer journeys but can instead be solved using other data sources. In addition to data resources discussed in this Report (Cross-platform panels, Streaming Only panels, Smart TV/ACR data and Federated 1P data), below are use cases and vendors that have an important role in solving for the full scope of content measurement use cases.

Content Use Cases Beyond Viewer Journeys

Demand Analytics

Forecasting consumers' appetite for content involves identifying the 'white space', the demand for content that is unserved or under-served by existing programming. What programming are consumers interested in that is not available in the market today? What programming is generating the greatest 'buzz' and conversation? Is there programming in the library that should be resurfaced and marketed to meet the demand? These are some of the data resources that support Demand Analytics:

- **Social media** - Social media platforms such as Meta, X, Snap and others are sources of consumer chatter and sentiment about films and video programming. To the extent that this buzz happens prior to release, social media is a core data source used by analytics providers to predict popularity and forecast revenue.
- **Apps delivering consumer benefits in exchange for data** - Companies such as Plum Research, Whip Media and Younify create apps for consumers to conveniently track and organize programming of interest and participate in participate in relevant discussion forums. In return for these benefits, consumers share their consumption data. Fan discussions may provide important clues for predicting content interest and potentially revenue.
- **"Piracy Networks"** - Piracy can be a leading indicator of program demand - what consumers want to watch, if they don't have budget or access constraints. Companies such as Muso and Cinelytic track pirated downloads and streaming through peer-to-peer and piracy websites to understand volume of pirated consumption, and consumption by title as a leading indicator of interest and demand.
- **Combined data sources to forecast demand.** Many vendors use a combination of data sources to obtain 'signal' for consumer demand for content, including Marketcast, Luminate Data, Parrot Analytics, ScreenEngine/ASI and Whip Media.

Content Valuation

Content valuation involves assigning a return on investment to content, aggregating information about content metadata, audiences, revenue as well as distribution strategies to estimate the profitability of each content and strategy. Today, publishers are using their own first-party streaming consumption data to measure content and supplementing it with second-party streaming consumption data provided by their distribution partners.

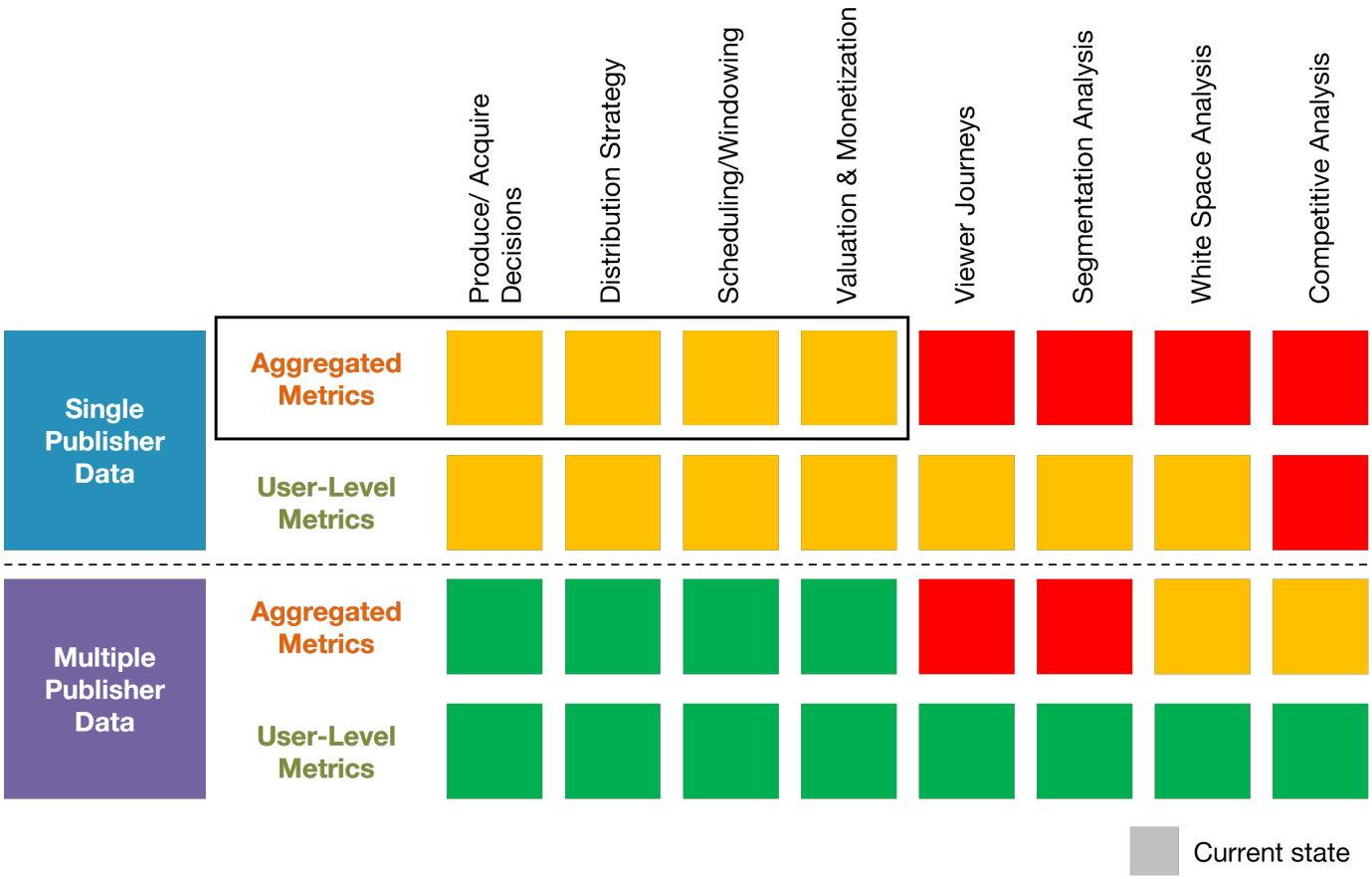
These data are typically combined and aggregated to create metrics, for example, about consumption in minutes or seconds for each title and distribution strategy. The data are further combined with revenue data from sources such as license fees, subscriber fees and advertising.

These data help answer questions such as: How does performance of one title compare to another? Is consumption growing over time? Which distribution strategy yields the largest audience? What is the value of the content to the distributor (and should we charge higher license fees?) This combined data is used as input for determining value to establish licensing fees and to make scheduling and windowing decisions. Companies such as Management Science Associates (MSA) process incoming first-, second- and third-party data by normalizing the data to enable comparisons across the various data sets.

The data are currently limited, however, in that each publisher can only view their own first-party data and, if provided, second-party data for their content from distribution partners and streaming content licensees. The data are siloed, lacking the benefit of following consumers on their journeys from one source of content to the next. The ideal measurement state for assessing content value would include user-based metrics for tracking viewer journeys across multiple publishers.

The table below illustrates how coverage of use cases improves in different data scenarios, based on whether aggregated or user-level data is available and whether the data is available for single publishers or multiple publishers. The shaded area in shows that the current state of content metrics that are largely aggregated and confined to single publishers, so missing the viewer journeys use case and other desirable use cases such as competitive and segmentation analysis. Once data from multiple publishers is in the mix, more use cases are covered. Multi-publisher user-level data covers use cases more comprehensively.

Data Collaboration Scenarios



Discovery

Discovery includes companies that provide EPG (electronic program guide) on consumer television sets, UX software that manages the user's experience in the program guide and on the TV, and companies that provide consumer apps to facilitate the consumer's discovery of content. Vendors include CCR Media, Comcast, Think Analytics, TiVo, Utelly, Whip Media and Younify.

Discovery data offers important insights into content development, content acquisition, and viewer acquisition and retention. There exists the potential to use these sources in tandem with third-party measurement to support content strategies.

Metadata & Schedules/Windows

Metadata refers to the descriptors for TV/video content, including series, episode, title, release timing, availability, talent, producers/directors, genre, mood and other descriptors. Schedules/windows refers to the capture of data about when and where the content ran. Both metadata and data about schedules/windows is used extensively by companies doing competitive analysis, forecasting demand and conducting valuations in order to more fully dimensionalize opportunities. Metadata and Schedules/Windows data vendors include Gracenote, IMDb, Stream Metrics, TiVo and others listed in the CIMMscape.

APPENDIX I: A Deeper Dive into Data

	Cross-Platform Panels (e.g., Nielsen, Hyphametrics)	**Streaming Only Panels** (e.g. YouGov, Digital i)	**Smart TV/ ACR Data** (e.g., Samba TV, Inscape)	**STB Data** (e.g., Comcast, Charter)	**Federated First- & Second-Party Data**
Representation	Not a gap; however, panelist effort to participate impacts response and cooperation rates.	Representation can be addressed with more rigorous methods. Today, Streaming Only panels are largely non-probability convenience samples.	Representation across source OEM/TV set brands is partially addressed using weighting and calibration, but groups such as non-connected TV users and OTA-only households are not measured.	Representation across source MVPDs is partially addressed using weighting and calibration, but groups such as streaming-only consumers and OTA-only households are not measured.	Not a gap. 1P and 2P data is "census" of streaming consumption.
Scale	Scale is a challenge due to panelist effort & privacy concerns, and high costs.	Scale is a challenge, but Streaming Only panels require lower panelist effort (compared to currency-grade panels) which reduces costs.	Not a gap, Smart TV data is from millions of Smart TV sets and households.	Not a gap, STB data is from millions of TV sets and households.	Not a gap. 1P and 2P data is a census of streaming consumption.
FAST	FAST can be measured with additional ACR/ fingerprinting.	Some providers cannot measure FAST channels that do not require user accounts. Other providers measure FAST. FAST and live streaming require additional data collection technology.	FAST OTT can be measured with additional ACR/ fingerprinting. Native app measurement - including FAST– is generally prohibited by streaming app owners.	FAST and other streaming is not measured. STB is linear TV only.	Not a gap for participating streaming app owners. FAST by non-participating app owners is not included.
Streaming	Expanding streaming titles requires additional ACR (however, business agreements may still prohibit reporting by some streaming app owners).	Not a gap, but providers need to expand beyond top streaming services and capture multiple streaming services from the same user. Virtually all publishers and titles can be measured.	If streaming is measured, it's OTT only today. Native app streaming requires a federated approach.	Streaming is not measured.	Not a gap for participating streaming app owners. Streaming by non-participating app owners is not included.

APPENDIX I: A Deeper Dive into Data

	Cross-Platform Panels (e.g., Nielsen, Hyphametrics)	**Streaming Only Panels** (e.g. YouGov, Digital i)	**Smart TV/ ACR Data** (e.g., Samba TV, Inscape)	**STB Data** (e.g., Comcast, Charter)	**Federated First- & Second-Party Data**
Mobile/PC	Mobile/PC consumption is measured via router meter. Expanding ACR/ fingerprinting is required to report titles.	Not a gap. Streaming Only panels capture consumption regardless of device.	Mobile/PC consumption requires streaming only panel, streaming app owner ACR permissions or streaming app owner 1P data.	Mobile/PC consumption is not measured. TV sets with an MVPD subscription only.	Not a gap for participating streaming app owners. Mobile/ PC by non-participating app owners is not included.
Linear TV	Not a gap. Measures all/ almost all linear TV.	Linear TV is not measured and would require an approach such as match with or recruitment via Smart TV or STB big data.	Not a gap, but only measures specific Smart TV brands and not all TV sets in the household.	Not a gap, but only measures TV sets from specific MVPD(s). Typically all sets in household.	Linear TV is not measured in 1P/2P data.
People	Not a gap. Measures people.	Measures accounts, not people. Accounts can be shared. Panelists could be required to set up individual accounts, but that's not included today. Co-viewing is also needed.	Measures Smart TVs, not people. To measure people requires co-viewing and a rigorous identity solution.	Measures TV sets**,** not people. Measuring people requires co-viewing and a rigorous identity solution.	Measures accounts and devices, not people. Meeasuring people requires co-viewing and a rigorous identity solution.

GLOSSARY

AI-based Content and Ads Recognition is an alternative to ACR/fingerprinting and watermarking that uses text, object tracking and navigation across viewer journeys (source, platform, program, content) within an AI framework to identify content and advertising. One advantage of this approach compared to watermarking is that it does not require the publisher's and marketer's cooperation to embed the watermark. One advantage of this approach compared to ACR/fingerprinting is that content and ads do not have to be provided in advance for the most accurate results. While the brand / brand variant in the ad can be identified using AI, the specific creative is identified with further, non-AI processes.

Automated Content Recognition (ACR) or fingerprinting is the process of periodically (every few seconds) creating a digital representation of the video or audio played on the TV set, and matching this digital representation back to a library of content to identify what was played.

TV set manufacturers (OEMs) and their partners which collect video consumption data from Smart TVs (such as Vizio, Samsung, Samba TV) and panel providers (such as TVision) use ACR as the technology to capture video consumption data. ACR technology is embedded in the TV's chipset or enabled via external devices.

ACR is the most commonly used technology to collect video consumption data because for linear TV consumption and streaming consumption on mobile and desktop devices, it does not require the cooperation of publishers or agencies/marketers to enable data collection. (Permission from streaming app owners is required for an OEM to capture streaming consumption data from Smart TVs). The disadvantage is that the best results occur when publishers and agencies/marketers submit their content, as well as associated metadata, for fingerprinting in advance, and they may choose not to do so. It can become very expensive to establish content libraries with fingerprints, especially where the content or ad is not provided in advance. Live programming, such as sports and news, cannot be fingerprinted in advance and is instead fingerprinted live. Typically, the most commonly watched content is fingerprinted (excluding some long tail content in libraries).

Another challenge is accurately identifying the source of the content because multiple distribution channels may air the same content, potentially at the same time, meaning the fingerprints (and associated content recognitions) would be the same. Providers may use location data, algorithms or on-screen identifiers (in the case of video fingerprinting) to identify the distribution source.

There are many technical details that are helpful to understand ACR and its applications for measurement that are beyond the scope of this Report, including the uniqueness of the fingerprints and the frequency of the fingerprints to accurately identify the content and the duration of consumption events.

Big Data in this Report refers to data about linear TV consumption collected from MVPD Set Top Boxes (STB) and Smart TVs ("Smart TV/ACR" data).

Broadband Only (BBO) refers to devices and/or households that consume video content only via Internet-connected devices and not Over the Air or via an MVPD subscription.

Content Measurement SDKs are software installed in the video player/server side that captures data beyond the basic app-level usage data available via router meters or mobile apps, so the data includes program identification.

Cross-platform panels refer to consumer research panels that are designed to be representative of the entire viewing population and measure both linear TV and streaming from the same consumers/panelists. Cross-platform panels are used for currency services and are typically recruited using probability sampling. Typically, households are recruited to install meters on their television sets and on their home Wi-Fi router to enable a measurement provider to capture video consumption data. In the U.S., Nielsen's meter is referred to as a People Meter. It uses both ACR/fingerprinting and watermarking technology. Panel households typically use remote-control devices to indicate who is viewing so that the service provider can assign demography. Data from cross-platform panels are often supplemented with first-party publisher data from logs, pixels or SDKs. There is now AI-based technology that can be used instead of ACR/fingerprinting and watermarking and Bluetooth technology to capture demography instead of remote-control devices, however, these technologies are not widely-deployed.

Direct-to-Consumer (DTC) refers to products and services that are marketed and sold directly to the consumer without an intermediary such as a retailer.

Federated Data Collaboration refers to a framework for governance, technology, data security and privacy to enable access to streaming app owner first-party data by other parties. In the context of content measurement this entails streaming app owners making their first-party streaming consumption data available for measurement by other parties using technologies that ensure data security and privacy such as data clean rooms.

GLOSSARY

First-party data in this Report refers to streaming consumption data generated when a consumer accesses a publisher's owned streaming app. For example, publisher NBCU owns the Peacock streaming app and the first-party data generated when consumers access the content.

Note that a publisher that owns streaming content may distribute or license the content to another streaming app owner. In this case, the other streaming app owner owns the first-party data. The other app owner's first-party data is 'second-party' data to the publisher that owns the streaming content. For example, NBCU licenses streaming content to distribution partner Tubi. Tubi is the owner of the first-party data generated when a consumer accesses the NBCU content through the Tubi app.

Whether a publisher has first-party streaming consumption data depends upon whether they own the streaming app. Whether a publisher has second-party streaming consumption data depends upon whether they have negotiated the rights to the second-party data with the streaming app owner, content distributor or other licensee.

Free Ad Supported Television (FAST) refers to pre-programmed streaming channels that are available to consumers at no charge. They are usually delivered via channels of pre-programmed content such as Vizio WatchFree+, Samsung TV Plus and LG Channels and are available in various apps including Amazon Freevee, Crackle, Pluto TV, Roku Channel, Tubi, Vudu and Xumo.

Linear TV refers to video consumption that occurs via a television set where the content is delivered via an MVPD subscription or Over-the-Air (OTA). Linear TV can include pre-recorded content, "live" content, video on demand (VOD) accessed via an MVPD and time-shifted consumption using a DVR. Linear TV also includes video delivered via a virtual MVPD (vMVPD) even though vMVPD is delivered to the consumer via the Internet. vMVPD is typically counted as Linear TV because the content available and user experience is similar/ the same as MVPD. Linear TV is distinguished from streaming in this Report.

Live Streaming refers to programming that is not pre-recorded and delivered via streaming. Live streaming examples includes sports programming and special events. FAST channels are similar to live streaming but FAST is pre-recorded programming.

Logs or publisher first-party log data is data captured on the publisher's or ad network's server that represents user request and/or delivery of content, including advertising. These data are often shared in aggregated form or in anonymized form (or via data clean rooms for privacy compliance) with marketers and their agencies to represent online ad impressions delivered.

Multi-Video Programming Distributor (MVPD) refers to video programming provided via a cable, satellite or fiber subscription service, such as Comcast, Charter, Dish, DirecTV or Verizon FIOS.

Native App refers to streaming video consumption that occurs via the publisher's app installed on a Smart TV or mobile/PC device and not consumption where the publisher's app is hosted by an intermediary OTT service. This Report distinguishes Native App consumption from OTT consumption because Native app streaming consumption data from Smart TVs cannot be collected by OEMs due to business agreements that prohibit such collection by the publishers. Not all publishers prohibit native app streaming data collection, however, the largest publishers do.

Original Equipment Manufacturer (OEM) refers to television set manufacturers, such as Vizio, Samsung, LG, Sony and others.

Over the Air (OTA) refers to video consumption that occurs via an antenna and not via an MVPD subscription or streaming.

Over the Air Only (OTA-Only) refers to households that consume video on television sets only via an antenna and do not use an MVPD subscription or streaming.

Over the Top (OTT) is often used generically to refer to any video consumption that occurs without an MVPD subscription. However, for the purposes of this Report, OTT refers to streaming consumption through a device and/or software external to the TV set, such as Apple TV+, Amazon Fire TV Stick, Roku and Chromecast. (Google discontinued support for Chromecast in 2023). This Report distinguishes OTT from Native App consumption because OTT streaming data can be collected from OEMs without violating their business agreements with publishers. Not all OEMs collect OTT streaming data, however.

People Meter refers to hardware and software that is installed in the TV set, attached via a dongle/USB port to the TV set or used in a portable form, to collect viewership and persons identification data. People Meters are the technology container typically using fingerprinting, watermarking, other technologies or combinations of technologies as the data collection methods. The People Meter technology also includes persons identification in some form, including:

- A remote control device that household members use to indicate who is viewing.

- Camera/sensor and image recognition attached to the TV set.
- A wearable device such as a wristwatch or mobile phone that is associated with a specific household member and has a Wi-Fi or Bluetooth communication with the meter.
- A personal device associated with a specific household member that detects watermarks or fingerprints based on the audio signal.

Once the fingerprint is generated or the watermark is recognized by the meter, it is timestamped and stored temporarily on the meter/data collection device. Since storage capacity is limited, it's important for data transmission to occur frequently (no less than daily). The technology transmits fingerprints and watermarks back to the measurement vendor typically using the home's Wi-Fi/internet connection. If the household does not have Wi-Fi/internet, the measurement vendor uses its own internet account for the purpose of data transmission. Nielsen's PPM system transmits via a device that uses the home's electrical wiring.

Probability Sampling refers to sampling techniques used to recruit consumer research panelists, where each person in the population has an equal and non-zero chance of being selected. Probability sampling is often distinguished from non-probability or "convenience" sampling, where panelists are recruited until the target number of participants is reached. Because probability samples can be projected to the population and have measures of statistical reliability, it is the preferred sampling methodology for currency-grade panels.

Router Meter is a stand-alone meter or software installed on the household's Internet router which monitors home network and internet traffic by source and device in the household. Router meters do not identify content title/name using this method, but it is possible to identify apps and devices that are used and the duration of consumption.

Second-party data in this Report refers to streaming consumption data generated when consumers access a streaming app owned by a publisher's distribution partner or streaming content licensee. For example, NBCU licenses streaming content to distribution partner Tubi. Tubi is the owner of the first-party data generated by consumer access to the Tubi app. The Tubi data is "second-party" data to NBCU.

Whether a publisher such as NBCU has second-party streaming consumption data from another streaming app such as Tubi depends upon whether they have negotiated the rights to the second-party data with the streaming app owner or other content licensee.

Set Top Box (STB) refers to Set Top Box, including the hardware and/or software provided to subscribers by MVPDs (multi-video programming distributor (e.g., Comcast, DISH)) that enables distribution of video content.

Smart TV/ACR data refers to data that is captured using ACR/fingerprinting software embedded in broadband Internet- connected Smart TV sets. It is typically comprised of a series of digital identifiers ("fingerprints") that are created and timestamped as the video or audio content plays on the TV set. See the above entry on ACR for more details about this data collection method.

STB data refers to data captured via return-path capabilities in the set-top box. The STB collects channel events and timestamps (such as whether the viewing events are via live feed, DVR or VOD, the household's interactions with the remote-control device, such as channel tuning interactions, on/off events, volume/mute events etc. Vendors then curate the data to match it to program titles and to convert events into start and stop times to measure duration.

Streaming refers to the continuous delivery of video content to the consumer via an app on a Smart TV, mobile/PC device or via an OTT service. Streaming is distinguished from downloads where the video content is physically stored on the device before it is consumed. Depending upon the measurement provider, streaming consumption may include downloads, typically based on an assumption that downloaded content is always viewed.

Tags (pixel tags, conversion pixel tags or tracking pixels) are software installed by marketers on their websites to capture website engagement and online purchase behavior by website visitors and to enable communication of these data back to publisher and agency web servers.

vMVPD refers to a "virtual" MVPD such as YouTube TV or Hulu + Live that provides cable, network and local programming on a subscription basis, much like traditional Pay TV services from cable, telecom and satellite providers. vMVPD subscribing households are often grouped with MVPD subscribing households for analysis purposes.

Video On Demand (VOD) refers to pre-recorded programming and includes both Ad-supported (AVOD) and subscription (SVOD). VOD is distinguished from Live streaming and FAST in this Report.

Watermarking is the process by which content providers (publishers for programming; marketers/ agencies for ads) embed a unique code in the content

which is periodically (every few seconds) read and timestamped by a meter or other device. The unique code is matched back to a library of content/ads and codes.

Watermarking has the advantage of requiring less data processing resources and potentially providing greater accuracy than fingerprinting. The library of watermarks is established in advance. Publishers and agencies/ marketers embed the watermarks in advance and/or at the time of distribution. There are fewer, if any, issues establishing distribution source.

However, the disadvantage is that it requires the cooperation of publishers and agencies/marketers to enable measurement, and they may choose not to do so. Watermarking is the main technology used by Nielsen which achieves a high level of cooperation across linear TV publishers, agencies and marketers. Most large, digital-only publishers have not historically cooperated with watermarking systems. Furthermore, some distribution methods can strip certain watermarks.


cimm
Coalition for Innovative
Media Measurement
ARF