

Global Cross-Media Data Requirements

Data Content for Cross-Media Integrations & APIs

November 2020

Background

Cross-media measurement implementations can vary depending on advertiser and local market needs. However, technical infrastructure for any cross-media measurement system should support a standard set of global data inputs that facilitate multiple reporting metrics, above and below standard thresholds (like MRC standard for video viewability), based on advertiser needs. This standardized, global data input enables consistency and viability across participating data providers as well as consistency in reporting for global advertisers.

Please note that the framework of this document and suggested global data inputs is not yet finalized and will likely vary once technical details for any given system are fully established. In addition, the proposed solution may be subject to terms between data providers and measurement service operators, including terms that may be necessary to comply with applicable privacy laws. However this document is an initial recommendation on these global data inputs which, once finalized, can support implementation work at the local level.

Several of the data fields enumerated below will require standardization related to formatting, definition and taxonomy. It is expected that future global and local workstreams will address the need for such standardization including, but not limited to, discrete metrics definitions (where not already present in industry guidance), Cross-Media campaign taxonomies and data collection and formatting guidance such as specific time thresholds for duration and playback accumulation. In addition, while this document describes high level requirements for the inputs and outputs of the system, as an immediate next step to the publication of this document a more detailed specification & schema of how impressions and these data fields should be encoded, as well as how these requirements are implemented throughout different APIs of the system will be published and open sourced as part of the Inputs and Outputs workstreams led by the WFA, ANA and ISBA.

Phases of Data Processing

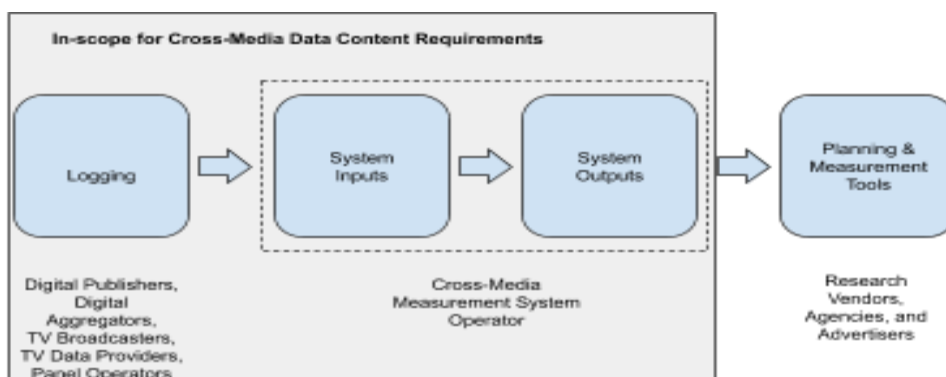


Fig 1. Phases of data processing for cross-media measurement

Cross-media measurement implementations typically have 3 phases of data processing:

- 1. Data Logging / Data Preparation:** Raw data that publishers, broadcasters, aggregators and other data providers collect, log, and process. This data is often collected by first-parties and can be processed privately to

formulate *Measurement system inputs*. There is a tradeoff between collecting/logging large amounts of 1st party data and ensuring high quality user experience on apps and/or websites (Heavy logging can significantly degrade app/website performance), so this tradeoff should be considered when finalizing the set of *Logging* data inputs.

- a. **Examples of Logging** = Impressions served, Impressions viewed by a human, pixels on screen, number of seconds viewed, etc.
2. **Measurement System Inputs:** Data inputs into the cross-media measurement service/system operator. This data is what is required of participating data providers to contribute into the cross-media measurement system. *Measurement system inputs* undergo data processing within the cross-media measurement system operator as per local cross-media measurement requirements to produce *Measurement system outputs*.
 - a. **Examples of Measurement System Inputs** = 100% pixels on screen flag, 2+ seconds video viewed flag
3. **Measurement System Outputs:** Data outputs of the cross-media measurement service/system operator for raw reports and APIs. This data is the output of the cross-media measurement service that enables agencies, research vendors, and advertisers to build new or integrate into existing cross-media measurement and planning tools.
 - a. **Examples of Measurement System Outputs** = Total impressions that are MRC-viewable (video).

“Delivered Impression” as the Minimum Threshold

It is important to define the minimum threshold that impressions should meet so as to qualify to be a *system input* into the cross-media measurement system. Including all impressions (even those that are served but not delivered or unrendered such as blocked ads) puts too much burden on the cross-media system to process and filter a lot of data, but restricting system inputs to only include comparison metrics (like MRC standards for video viewability) remove a lot of flexibility in analysis that advertisers have requested for in cross-media measurement. Hence, it is preferable to include only “IAB/MRC-defined/standardized delivered or rendered impressions” as system inputs, which allows local cross-media implementations to create various reporting metrics including, but not limited to, standard comparison metrics like MRC viewability standards.

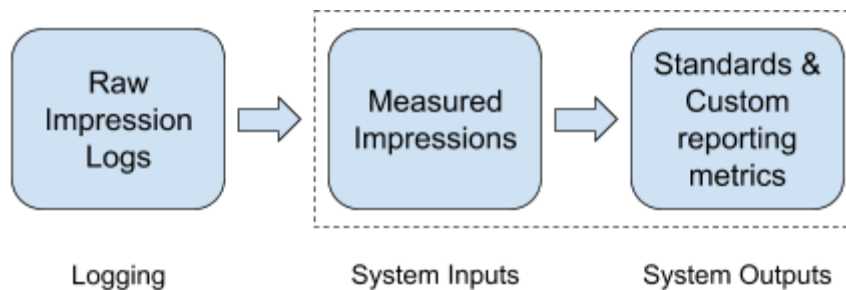


Fig 2. Impression thresholds across various phases of data

Detailed Data Requirements

The following provides possible data fields that would create the base for the Global Cross-Media Measurement infrastructure. The categories are 'builds' dependent on the desired metric category. Data fields can be supplied by

different vendors and integrated into the overall data-set. Broader or more conclusion-oriented metrics (like “viewable Y/N”) may be accepted as inputs for data provided by independent third-party measurement entities that have been audited and accredited (for example, accepting viewable yes/no, rather than the time and pixel state details), more on this below (see Audits section).

These requirements, outlined in the table below, are not meant to be prescriptive and are subject to further modification and implementation at local market levels. Additional evaluation will be required through the lens of the technical blueprint, to ensure required data parameters still allow the measurement system to operate efficiently with respect to data processing and cost. Additionally, evaluation will also be required to ensure that user privacy guardrails are fully upheld.

Fields have been segmented by those required to be collected, maintained and processed by first-parties or aggregators (“Logging”), those that should be provided to the measurement system (“Measurement System Inputs”), and those expected as outputs of the measurement system (“Measurement System Outputs”). *Measurement System Outputs* fall into two categories - metrics (these are the actual values that get calculated and surfaced in dashboards) vs. data cuts (these are the filters and group-bys surfaced in dashboards).

Note:

1. Some data fields only need to be “logged”, whereas others also need to be shared as “Measurement System Inputs” to the Cross-Media Measurement Service Operator. Similarly, some data fields need to be available across all 3 phases of data processing. The Y/N entries in the table below signify whether a particular data field is required in that phase of cross-media measurement data processing.
2. Data fields that are shared as “Measurement System Inputs” for all impressions will not be shared as raw data, but will go through additional processing to transform them into privacy-preserving data structures such as Bloom Filters.

Category	Data	Logging	System Inputs	System Outputs		Notes
				Metrics	Data Cuts	
01 General	Platform (OTT, TV, Desktop, Mobile Web, App)	Y	Y		Y	
	Browser	Y	Y		Y	
	Device Type	Y	Y		Y	
	Device ID	Y	N	N	N	Used for VID
02 Ad Impressions (Digital & TV)	Unique Impression ID	Y	N	N	N	To aggregate campaigns, creatives
	Render Timestamp	Y	N	N	N	To derive data date, daypart

Category	Data	Logging	System Inputs	System Outputs		Notes
				Metrics	Data Cuts	
02 Ad Impressions (Digital & TV) continued...	Render (Y/N)	Y	Y	N	N	To derive rendered impressions
	Rendered Impression	Y	Y	Y		
	Data Date & Daypart	Y	Y		Y	
	Advertiser ID	Y	Y		Y	
	Campaign ID	Y	Y		Y	
	Unique Creative ID	Y	Y		Y	
	Creative Type	Y	Y		Y	
	Broadcaster / Publisher ID	Y	Y		Y	
	Channel / Site / App ID	Y	Y		Y	
	Placement ID	Y	Y		Y	
	URL (Domain Name)	Y	Y		Y	
2A Viewability	Measurement Source (Tag, SDK, VPAID, MRAID, etc.,)	Y	N	N	N	
	Pixels	Y	N	N	N	To derive counts / time vs. predefined pixel thresholds
	Measurable Indicator	Y	Y	N	Y	To derive measured impressions
	Total and Max Time (thresholds tbc) at Predefined Pixel Thresholds (0%, 50%, 100%)	Y	Y	N	Y	To derive viewable impressions
	Viewable Y/N	Y	Y	N	Y	To derive viewable impressions
	Strong User Interaction	Y	Y	N	N	To derive viewable impressions
	Measured Impression	Y	Y	Y		
	Viewable/Non-viewable Impression	Y	Y	Y		
2B Digital Display/Rich Media	Auto Refresh Y/N	Y	N	N	N	

Category	Data	Logging	System Inputs	System Outputs		Notes
				Metrics	Data Cuts	
2B Digital Display/Rich Media continued...	Creative Size	Y	Y		Y	
	Rich Media State (expanded / unexpanded)	Y	Y		Y	
	Default/Backup Image Detail and Use	Y	Y	N	Y	
2C Video Ads (Including OTT)	Playback Time / Duration (total or threshold based)	Y	Y	Y		
	Viewable Duration (total or threshold based)	Y	Y	Y		
	Quartile Complete Yes/No Q1-Q4	Y	Y		Y	
	Auto-Play/Click to Play	Y	Y		Y	
	Forced Duration	Y	Y		Y	
	Audio Status (Mute/Non-zero, Captions)	Y	Y		Y	
	Creative Size	Y	Y		Y	
	Creative Length	Y	Y		Y	
	Completion (Y/N)	Y	Y		Y	
03 Digital Video Content (Including OTT)	Content Identifier	Y	Y	N	N	To derive content segments / title
	Continuous Play	Y	Y	N	Y	
	Audio Status (Mute, Non-zero, Captions, Unknown)	Y	Y		Y	
	Ad Insertion Model Dynamic/Static	Y	Y		Y	
	Ad Supported Yes/No	Y	Y		Y	
	Monetizable Yes/No	Y	Y		Y	
	Content Segment / Title	Y	Y		Y	
	Playback Time & Duration	Y	Y	Y		
	Viewable Duration	Y	Y	Y		
	Content views	Y	Y	Y		

Category	Data	Logging	System Inputs	System Outputs		Notes
				Metrics	Data Cuts	
04 Television Content	Uncredited Tuning Contributing	Y	N	N	N	To derive playback time
	Peripheral Device Info	Y	N	N	N	
	Headend	Y	Y	N	N	
	HD/SD	Y	Y	N	N	
	Source/Channel	Y	Y	N	N	To derive network / local affiliate
	Broadcast Time/Viewing Time (live or TSV)	Y	Y	N	N	To derive data date, daypart
	Audio Status (Mute, Non-zero, Captions, Unknown)	Y	Y	N	Y	
	Viewing Location (In-Home/OOH)	Y	Y	N	Y	
	MVPD/vMVPD/Operator	Y	Y	N	Y	
	Ad Supported Yes/No	Y	Y		Y	
	Ad Insertion Model Dynamic/Static	Y	Y		Y	
	Network or Local Affiliate	Y	Y		Y	
	Program ID/Name	Y	Y		Y	
	Data Date & Daypart	Y	Y		Y	
	Playback Time & Duration	Y	Y	Y		
	Viewable Duration	Y	Y	Y		
05 Invalid Traffic Filtration (IVT)	IP Address (X-Forwarded-For/MAC)	Y	N	N	N	To derive GIVT and/or SIVT
	User Agent	Y	N	N	N	To derive GIVT and/or SIVT
	Cookie, Unique ID, Mobile Fingerprint	Y	N	N	N	To derive GIVT and/or SIVT
	Carrier Information	Y	N	N	N	To derive GIVT and/or SIVT
	Location	Y	N	N	N	To derive GIVT and/or SIVT

Category	Data	Logging	System Inputs	System Outputs		Notes
				Metrics	Data Cuts	
05 Invalid Traffic Filtration (IVT) continued...	Pre-Fetch Headers	Y	N	N	N	To derive GIVT and/or SIVT
	Open RTB Attributes	Y	N	N	N	To derive GIVT and/or SIVT
	App Identifier	Y	Y	N	N	To derive GIVT and/or SIVT
	General Valid/Invalid	Y	Y		Y	
	Sophisticated Valid/Invalid	Y	Y		Y	
06 Audience (Demos and Behaviors)	Granular Demo (Age, Gender, Race, Ethnicity, Language)	Y	N	N	N	To derive demo delineated metrics
	IP Address (X-Forwarded-For/MAC)	Y	N	N	N	To derive demo delineated metrics
	User Agent	Y	N	N	N	To derive demo delineated metrics
	Cookie, Unique ID, Mobile Fingerprint	Y	N	N	N	To derive demo delineated metrics
	Unique User ID	Y	N	N	N	To derive demo delineated metrics
	Data Origin (First/Third Party, Declared, Observed, Modelled)	Y	N	N	N	To derive demo delineated metrics
	Aggregated Demo (Age and Gender)	Y	Y	N	Y	To derive demo delineated metrics
	Reach	N	N	Y		Derived metric
	Frequency	N	N	Y		Derived metric
	Household Data (where applicable)	Y	Y	N	N	To inform TV metrics
	Additional aggregated Demo's (Race, Ethnicity, Language etc.)	Y	Y	N	Y	
	Presence Indicators	Y	Y	N	Y	
	Purchase History	Y	Y	N	Y	
Geography	Y	Y	N	Y		

Note: Additional categories for future consideration include brand safety, outcomes, and future formats.

Audits

To ensure consistency and accuracy, data processing audits will likely be required. Data providers that have 1st party mechanisms to log user data may only need to provide *System Inputs* into the measurement system. To ensure consistency across data providers, this might require 3rd party auditing/verification of *Logging* mechanisms as dictated by advertiser perspectives. These audits - in addition to audits on *System Inputs* - should be done at the global level for data providers that operate globally.

Data providers that do not have 1st party mechanisms to log user data can choose to work with independent 3rd party service providers (that will be audited) that perform the processing on behalf of the data provider and provide these Measurement Inputs into the cross-media system.

In addition to the global audits on Logging, measurement service operators will be required to enable 3rd party auditing/verification of *System Inputs* (for non-global data providers) as well as *System Outputs*. This will be done at the local level since measurement service operators could vary by market.

The specific details of the audits, including those for global data providers, are still to be determined and agreed upon.