

# Advertising Metadata Standardization Initiative

ARF TOWN HALL  
October 20, 2020

# Outline

What is AMSI?

The Challenge

A Problem Worth Solving

Approach

Discovery

Solution

Adoption

Call to Action

# Presenters



**Jane Clarke**

*CEO, Managing  
Director*

**CIMM**



**Vijoy Gopalakrishnan**

*SVP/Principal, Media  
Center of Excellence*

**IRI**



**Harold S. Geller**

*Executive Director*

**Advertising Digital  
Identification LLC  
(Ad-ID)**



**Jessica Hogue**

*General Manager  
of Measurement  
& Analytics*

**Innovid**

# What is AMSI?

---

- ▶ Collaboration between media industry associations and companies involved in digital and TV attribution/ROI analysis to standardize nomenclature, metadata fields and file formats throughout the ecosystem.

# Partners



comscore

geopath

group<sup>m</sup>

iab.

iab.  
TECH LAB

INNOVLD



IRi

iSpot.tv

KANTAR

LiveRamp®

nielsen  
.....

NinthDecimal

Standard  
Media  
index

# The Challenge

---

- The proliferation of media platforms has exposed marketers to new insights related to the efficacy and impact of their cross-media investments, but the collateral damage comes in the form of differing data formats and levels of extraction capabilities.
- There are no standards on ad exposure file formats within digital platforms and between digital and television, and no incentive to standardize data or help make it seamless.
- Marketers may have to deal with as many as 200 different data formats – quite a humbling experience.

*Source: Italicized Text (8 point)*

# Differing Formats and Fields from Digital AdEx Exposure Logs

e.g. Highlighted time stamp uses variations of exposure time, server time, receipt time, etc. in differing formats

Format 1	Format 2	Format 3	Format 4	Format 5	Format 6	Format 7
CUSTOMER_LINK	Customer_Link	Customer_Link	CUSTOMER_LINK	CUSTOMER_LINK	Customer_Link	CUSTOMER_LINK
CAMPAIGN_ID	field	exposure_date	EVENT_TIME	ID_METHOD	AdvertiserID	CAMPAIGNID
PLACEMENT_ID	Log_Date	campaign_id	ADVERTISER_ID	UNIX_TIMESTAMP	AccountID	BANNERID
IP	Log_Time	placement_id	CAMPAIGN_ID	PLATFORM_NAME	EventTypeID	BANNERSITEID
EVENT_TYPE	Log_Time_UTC_Se	creative_id	AD_ID	ORDER_ID	PlacementID	PLACEMENTID
RUN_TSTAMP	Ref_URL		RENDERING_ID	CAMPAIGN_ID	SiteID	CREATIVEID
	ID2		CREATIVE_VERSION	DEVICE_TYPE	CampaignID	HITDATE
	campaign_id		SITE_ID_DCM	DEVICE_OS	BrandID	
	creative_id		PLACEMENT_ID	DEVICE_MAKE	EventID	
	placement_id		COUNTRY_CODE	DEVICE_MODEL	InteractionID	
	site_id		STATEREGION	CARRIER	InteractionDuration	
	User_IP		BROWSERPLATFORM	BROWSER	VideoAssetID	
			BROWSERPLATFORM	PLACEMENT	InteractionDate	
			OPERATING_SYSTEM	MEDIA_TYPE	EntityID	
			DESIGNATED_MARKET	CREATIVE_ID	PCP	
			CITY_ID	ACTION		
			ZIPPOSTAL_CODE	TARGETING_TACTIC		

## Differing formats and fields

Format1	Format2	Format3	Format4	Format5	Format6	Format7
Customer_Link	Customer_Link	Customer_Link	Customer_Link	Customer_Link	Customer_Link	Customer_Link
campaign_id	BUNDLE	Event Time	Event Time	event_timestamp	exposure_date	ID Method
placement_id	CAMPAIGN_ID	Advertiser ID	User ID	advertiser_id	campaign_id	Unix Timestamp
ip	SITE_NAME	Campaign ID	Advertiser ID	campaign_id	publisher_id	Platform Name
event_type	PLACEMENT_ID	Ad ID	Campaign ID	placement_id	placement_id	Order ID
run_tstamp	CREATIVE_ID	Rendering ID	Ad ID	ad_id	creative_id	Campaign ID
		Creative Version	Rendering ID	video_id	var1	Device Type
		Site ID (DCM)	Creative Version	inventory_source_id	var2	Device OS
		Placement ID	Site ID (DCM)	data_provider	var3	Device Make
		Country Code	Placement ID		var4	Device Model
		State/Region	Country Code		var5	Carrier
		Browser/Platform ID	State/Region		var6	Browser
		Browser/Platform Version	Browser/Platform ID		var7	Placement

# A Problem Worth Solving

- This is a challenge worth solving to maximize the value of ad exposure data leading to:
  - Commercial benefit through automation
  - Increased speed and scalability
  - Seamless interoperability across the ecosystem
  - Transparency
  - More accurate cross-media measurement of outcomes
- For digital media, this will require a common standard across all the metadata normally ingested for a campaign among ad exchanges, servers, DMPs, DSPs and agencies.
- For TV, this also requires more standardized metadata, file formats and ad identifiers between network logs, ad verification services and third-party processors.



# Approach

---

## ▶ Built partnerships

- Recruited ecosystem companies and associations with “skin in the game.”

## ▶ Launched in February 2020

- CIMM 2/6/20 Summit
- AdExchanger article (<https://www.adexchanger.com/data-driven-thinking/why-we-need-a-upc-code-for-ads/>)

## ▶ Completed Discovery Process

- Factfinding mission with group partners: Dataroma, Fox, iSpot, Kantar, & Nielsen
- Drafted metadata standard and dataflow

## ▶ Developing Advocacy and Adoption Strategy



# Why We Need A UPC Code For Ads

by AdExchanger // Monday, September 23rd, 2019 – 12:05 am

*"Data-Driven Thinking" is written by members of the media community and contains fresh ideas on the digital revolution in media.*

*Today's column is written by Vijoy Gopalakrishnan, senior vice president and principal at the [IRI](#) Media Center of Excellence.*

Fifteen years ago, ad tech was a manual affair. Piecing together ad serving was such a headache.

While that headache may have virtually disappeared, the growth in ad servers, demand-side platforms (DSPs), data management platforms (DMPs), supply-side platforms and onboarders has created another challenge: the interpretation of data files.

The proliferation of platforms has exposed marketers to new insights, but the collateral damage comes in the form of differing levels of data formats and extraction capabilities. There are no standards on ad exposure file formats and no incentive for platforms to standardize data or help make it seamless.

Marketers may have to deal with as many as 200 different data formats – quite a humbling experience. The advertising industry has successfully created highly advanced analytical solutions, such as exotic machine learning algorithms and zero human touch, multithreaded, in-memory computing products, yet marketers are brought to their knees by lowly nonstandard file formats.

In speaking with industry veterans across agencies and ad tech, market research and CPG companies, this theme repeats itself. The first priority for all is to serve ads, making standardization of data formats a third or even fourth priority. In the heat of campaign execution, we heap casual cruelty on inputs, outputs and descriptor standardization.

Can we all take a collective breath and step back?

## A problem worth solving

This is a challenge worth solving to maximize the value of ad exposure data. What lies at the end of the tunnel is a commercial benefit through automation (bottom-line impact), increased speed and scalability (top-line impact), seamless interoperability across the ecosystem and transparency.



Vijoy Gopalakrishnan  
Principal, IRI Media  
Center Of Excellence

**Data-Driven  
Thinker**

The challenge is formidable but worth it. This will require a common standard across all the metadata normally ingested for a campaign among ad exchanges, servers, DMPs, DSPs and agencies.

## Progress being made

Several organizations are already making progress against this goal, including Ad-ID, Coalition for Innovative Media Measurement (CIMM), IAB Tech Lab and the Society of Motion Picture and Television Engineers (SMPTE). Ad-ID, an advertising asset identification system, has improved cross-platform adoption and continues to educate the industry on the importance of standardized metadata. The metadata is comprehensive, including ad title, length, product industry, etc., and goes a long way in meeting the needs of asset-related information.

CIMM, Kantar and SMPTE developed the Trackable Asset Cross-Platform Identification (TAXI Complete) initiative, which includes an open standard watermark embedded into audio TV or digital files that incorporate the Ad-ID code and associated metadata. TAXI Complete is new to the ecosystem and needs adoption and advocacy to make the watermark ubiquitous.

Another piece of this solution requires greater adoption of the updated VAST protocol developed by the IAB Tech Lab. It includes the Ad-ID code in the United States as part of the Universal Ad ID, which provides the framework for increased cross-platform interoperability. This, too, will only be successful if more vendors initiate adoption of VAST 4.

While much work has been done to advance standardizing metadata, there is still much more work to do. The IAB Tech Lab, CIMM, 4A's and other industry associations need to collaborate on expanding the standards for advertising metadata to reduce, or even better eliminate, the many variations of data formats currently in use. It is in the best interest of the entire ad tech ecosystem to prioritize standard metadata within the industry.

Follow IRI ([@iriworldwide](#)) and AdExchanger ([@adexchanger](#)) on Twitter.

# Discovery - Time Zone

Time zone and format is not fully standardized for Digital or TV, although TV has done more to standardize.

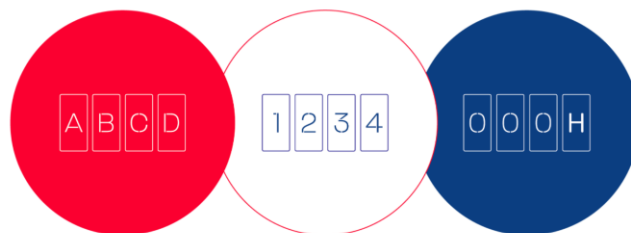
- TV data providers such as Kantar, iSpot and Nielsen use listening stations to audit ad clearance and use their listening station times to verify and standardize time and zone (usually output to Eastern time)
- All the above TV providers agree that a standard such as UTC with a time stamp convention (such as HH/MM/SS and either a string or number) would help reduce manual intervention
- Nielsen does receive logs from stations/networks and where there is a discrepancy the station logs supersedes other data. However, this is primarily to identify the minute in which the commercial ran (as input into the C3/C7 Ratings currency), and not to provide actual “ad audience measurement” yet.

# Discovery - Ad Creative and TV Program IDs

- TV Program data nomenclature are not standardized using the available EIDR standard (Entertainment ID Registry). Vendors use custom solutions or license program names from Nielsen/Gracenote.
- Ad Creative ID is not standardized despite the availability of a standardized identifier (Ad-ID). Even where standardized to Ad-ID, there are versions of the parent creative (with Ad-ID) that do not use an Ad-ID or mapping to the parent Ad-ID

# Ad-ID is the Industry Standard

- Unique alpha-numeric code assigned to a commercial asset when it is “born.”
- “Slate” connected to commercial containing metadata pertinent to the spot.
  - Using Ad-ID’s Complete External Access (CEA) API
- Managed by advertiser or agency through cloud-based platform.
- Permission based access across multiple agencies, ad tech and publishers.
- JV between ANA and 4A’s. (ISCI R.I.P. 2007)



# Ad-ID Adoption Stats

---

- Over 4,000 Advertisers Using Ad-ID Represent:
  - 84% of Network Dollars
  - 70% of Cable Dollars
  - 43% of Spot TV Dollars
  - 66% of All TV Dollars
- Ad-ID Is Used by Over 900 Agencies
- Over 2.5 Million Ads Registered

# Digital Media - Different IDs, Same Creative

- A single ad that should have been displayed 1 time is displayed multiple times because each entity (Agency, Publisher, DSP, SSP) have assigned proprietary identification.



# Benefits of Ad-ID

## IMPROVED POST-DELIVERY REPORTING

Knowing exactly where an ad ran and when is key to post-campaign wrap-ups. With more detailed information on the performance of creative asset, the better a new campaign can be created.

## IMPROVED REACH AND FREQUENCY

Using Ad-ID allows for the accurate delivery of creative assets. Accuracy ensures the right message gets delivered to the right consumer, in an appropriate number of times.

## ADDRESSES AD COLLISION

When multiple ad servers insert ads into the same video, they need to acknowledge a common identifier, or it is impossible at scale to prevent the same ad or advertiser from running back to back.

## REDUCED FRAGMENTATION ACROSS SCREENS

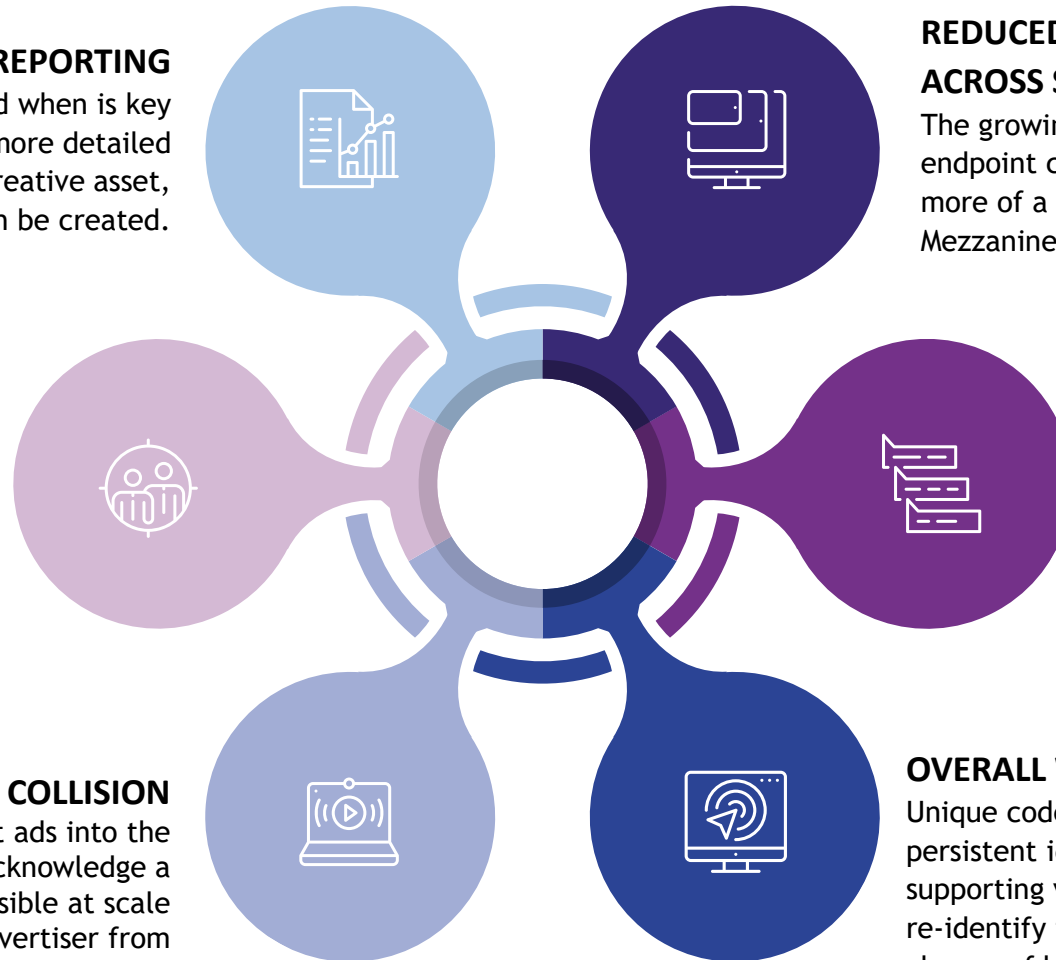
The growing landscape of screens and endpoint creative requirements caused more of a need for one master file - Mezzanine

## DEDUPLICATED STORAGE OF ADS

Demand and supply side vendors hosting and storing creatives can avoid unnecessary storage of duplicate source media files.

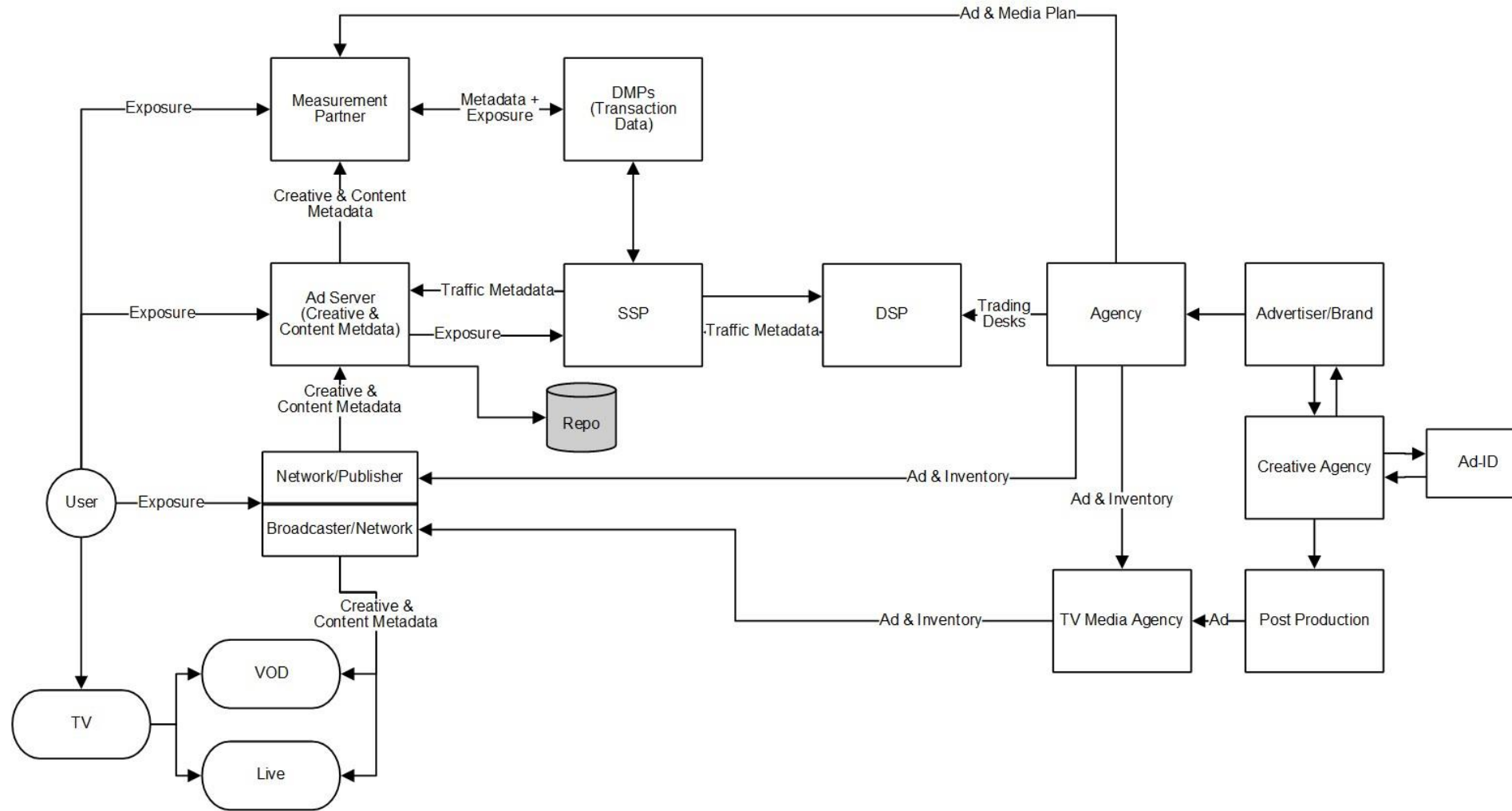
## OVERALL WORKFLOW EFFICIENCIES

Unique code generation enables a persistent id across all channels and supporting vendors, reducing the need to re-identify the asset or increase the chance of human error.





# Discovery - Ad Exposure Data Flow Digital/TV



# Solution

Required Ad Metadata elements for a standardized output solves the AMSI challenge

Ranking	Metadata Type	Video Only?	Example	Required?
1	Cross Device ID		alphanumeric value	Yes
1	Cross Device ID Type		ID Spaces (i.e. IdentityLink)	Yes
2	Timestamp		Unix timestamp	Yes
2	Timezone			Yes
3	Duration Played?		mm:ss:ms	Yes
3	Video Events	Yes	start, first_quartile, etc.	Yes
4	Session ID		numeric value	Yes
5	Device		phone, tablet, desktop	Yes
5	Device ID		Hashed IDFA/AAID/STB ID	Yes
6	IP Address		Hashed (i.e. SHA1)	Yes
7	Operating System		Android 8.0.2	Yes
8	Ad-ID		alphanumeric value (ad-id.org)	Yes
9	Creative (Ad Copy) ID		numeric value	Yes
10	Network/Publisher/App ID		numeric value	Yes
10	Domain/Network/App Name		ESPN.com/ESPN	Yes
11	Content Delivery Type		VOD, Live, audio stream, etc.	Yes
12	Content Genre		Comedy, Sports, Health, etc.	Yes
13	Program Title	Yes	The Big Bang Theory	Yes
14	Episode Title	Yes	The Conference Valuation (S12 E17)	Yes
15	Placement ID		numeric value	Yes
16	Creative Type		display, video, audio, etc.	Yes
16	Creative Length (sec.)	Yes	mm:ss:ms	Yes
17	Campaign ID		numeric value	Yes

Ranking refers to the order of operations from an ad consumption standpoint

# In Sum...

---

- Consistent use of Ad-ID to tag all creative versions
- Standardized Timestamp for exposure (UTC zone, HH/MM/SS number or string format)
- Standardize program description universally
- Consistent use of all the metadata listed across all exposure file transmission

...will help the industry achieve

- Transparency
- Seamless interoperability across the ecosystem
- Increased speed and scalability
- Reduced costs
- Improved accuracy in cross-media measurement of outcomes

# Adoption

---

- There is general agreement on the need for standardization; some DSPs have shown willingness to codify the standard in their products. Even Nielsen and Kantar agree that there could be some standardized metadata and the use of the Ad-ID between the two competing services.
- IAB Tech Lab and MRC are aligning recommended standards and data fields to create a multi-pronged influencing strategy and user friendly “data scapes” that are aligned with the AMSI initiative.
- Plan for outreach to digital publishers (Google, FB) and DSP/SSP needs to be planned out.

# Call to Action

---

- Adoption will reduce, or even better eliminate, the many variations of data formats currently in use, which will increase accuracy and efficiency and support automation and scaling.
- It's in the best interest of the entire ad tech ecosystem to prioritize standard metadata & ad identification within the industry.
- Contact us to learn how you can support AMSI!

# Panelists



**Travis Clinger**

*SVP, Addressability  
and Ecosystem*

**LiveRamp**



**Harold S. Geller**

*Executive Director*  
**Advertising Digital  
Identification LLC  
(Ad-ID)**



**Vijoy Gopalakrishnan**

*SVP/Principal, Media  
Center of Excellence*

**IRI**



**Jessica Hogue**

*General Manager  
of Measurement  
& Analytics*

**Innovid**



**Jane Clarke**

*CEO, Managing  
Director*

**CIMM**

**Moderator**

**Travis Clinger** [tclinger@liveramp.com](mailto:tclinger@liveramp.com)

**Harold S. Geller** [hgeller@ad-id.org](mailto:hgeller@ad-id.org)

**Vijoy Gopalakrishnan** [Vijoy.Gopalakrishnan@IRIworldwide.com](mailto:Vijoy.Gopalakrishnan@IRIworldwide.com)

**Jessica Hogue** [Jessica.Hogue@innovid.com](mailto:Jessica.Hogue@innovid.com)

**Jane Clarke** [Jane.Clarke@cimm-us.org](mailto:Jane.Clarke@cimm-us.org)

# THANK YOU