 **August 2016**

**RFP for Combining Smart TV and Set Top Box Data**

1. **Background & Objectives**

As part of its initiative to bring more granular audience measurement to the television industry, the Coalition for Innovative Media Measurement (CIMM) is interested in combining data from Smart TVs with that of Set-Top Boxes to create a nationally projectable “mega-sample” of TV tuning data. Both datasets complement each other and together can provide some of the data that the other lacks for the purpose of creating a nationally representative dataset of TV tuning. This proposal is originating with CIMM’s Smart TV/Advanced TV Committee, and Committee members would participate in the project in an ongoing consultative role.

Set-top box data are currently being made available to media measurement vendors (and also to audience based planning and buying platforms) from Dish, AT&T U-verse/Direct TV, Charter, Cox, TiVo, and Fourth Wall Media. However, these data only represent about 40% of TV tuning in the U.S. and particularly lack coverage of Cable TV MVPDs in major U.S. markets such as New York, Chicago, L.A., Boston and Philadelphia. Comcast, Time Warner Cable (now New Charter) and Cablevision (New York DMA only) are using household-level data for their internal ad sales purposes, but haven’t made the data available to external vendors or networks/content providers, except in the form of summary reports. This has created a need to find alternative sources of census-level TV tuning data from Smart TVs and to combine both datasets to create and/or more accurately model a nationally representative footprint.

Smart TVs have certain advantages over STBs, such as the ability to know when the TV set is on or off. In Pay TV environments (except for some AT&T U-Verse homes and TiVo), the TV set can be turned off while the STB is still on. This results in over-reports TV viewing with “false positive” tuning events. MVPDs and measurement firms have developed “cap and edit” rules for cutting off the tuning events, based on genre of show and time of day. Smart TV data could be used to refine the “cap and edit” rules.

Additionally, Smart TVs report all content that “hits the screen,” whereas STB data only reports linear viewing that passes through the STB or network (in the case of network monitoring). Smart TVs can report tuning to VOD and SVOD (from cable and OTT) and interactive games using the TV screen, even if they aren’t “linear TV.” The only challenges Smart TVs have in reporting content is the limitation of what content has been ingested into the reference library and to which the fingerprints are matched. Smart TVs have difficulty identifying the source for crediting the viewing, especially if the content is identical across source (i.e., same content on different channels or SVOD providers, with same ad load/same fingerprint).

Smart TVs can also report data in near “real-time,” although the ACR technology has errors in detection. These need to be corrected regularly and updated reporting needs to be produced on a rolling basis as corrections are received – typically within a few days. Conversely, STB data is typically reported on a weekly basis, primarily due to “messy” and incomplete/missing data that need to be cleaned, edited, matched to program names, weighted and statistically balanced to represent the national TV universe. Perhaps by combining the data, there might be a better solution for releasing tuning data in stages.

A strength of MVPDs is that they have a clear 1st party billing relationship with their households and are able to provide an accurate match to data enrichment providers, using third-party blind matching (privacy-compliant) procedures. Smart TVs can only be matched to their respective household addresses using IP address, which typically is dynamic (i.e., changing frequently). Smart TV ACR vendors have developed methods of using other data from apps on the Smart TVs (such as amazon) who do have 1st party billing relationships with households and also can link both devices and IP addresses within a home wi-fi network to enable a HH match to supply physical name/address (in a privacy-compliant manner). Perhaps MVPD data can help to supplement HH identifiers for Smart TVs, if it can be done in a privacy-compliant way.

For these reasons, and possibly for more reasons not yet anticipated, combining data for Smart TVs and STBs within the same households could aid in the development of modeling algorithms and new procedures that could be applied to increase the value of the combined datasets.

1. **Method & Deliverables**

We are open to considering a variety of approaches, either from research/technology vendors who are licensing both kinds of data, or from either STB or Smart TV providers who want to combine both datasets, to “Big Data” firms who could provide the infrastructure for combining large datasets.

At this point we’re assuming that we would need to conduct at least two forms of analyses:

1. ***Creation of nationally representative footprint:*** The footprints of RPD from each STB and Smart TV provider would be compared to determine how best to statistically combine them to create a nationally representative footprint for all kinds of viewing environments: cable, satellite, OTT, OTA (might still be lacking), broadband/OTA, broadband only, etc.
2. ***Overlap homes comparison:***  A reasonably-sized subset of homes would be identified that have both STB data and Smart TV data to help analyze the following factors:
	1. ***Refine Set On/Set Off cap and edit rules for STB data from Smart TV data***
	2. ***Develop algorithms from STB homes to model tuning data for unconnected sets in homes with sets other than Smart TV***
	3. ***Develop algorithms for modeling SVOD/games use from Smart TVs to STB-only HHs***
	4. ***Use of Smart TV data to improve reporting of time-shifted tuning events in STB HHs***
	5. ***Validate privacy-compliant ways to match IP address to HH name/address***
	6. ***Use both datasets to create rolling episode tuning report from Day 1 to Day 35***
3. ***Additional considerations:*** Below are a few additional considerations that CIMM’s Smart TV/Advanced TV Committee would like to be addressed in your proposal:
* Can your methodology be transparent for the end users and ultimately audited by the MRC?
* How can new data sources (either STB or Smart TV data) be added, as they become available?
* Can the data be offered both as a syndicated product (with an online interface) and for custom analyses?
* Can the new dataset be integrated in a privacy compliant way with other datasets (at the household level), such as purchasing or other behavioral or proprietary datasets or digital ad tags or other cross-media usage data or panel data, for the purposes of targeting, cross-media optimization and/or to demonstrate post-exposure purchase behavior?
* Are there other issues you would like to raise and/or what additional analyses would you like to conduct?
1. **Proposal Schedule & Budget**

Written proposals/budgets are to be submitted by September 30, 2016. In the proposal, feel free to outline several approaches to address the challenges, each with their own cost and timing considerations, so that the Committee will have options to evaluate. Note that Committee members can also be available for interim meetings/calls to review and discuss potential approaches.

All submissions should be sent to:

Jane Clarke

CEO, Managing Director



Coalition for Innovative Media Measurement

1115 Broadway, 12th Floor

New York, N.Y. 10010

212-710-2638

jane.clarke@cimm-us.org