

REPORT TO THE  
SET-TOP BOX COMMITTEE  
OF THE COUNCIL FOR RESEARCH EXCELLENCE



---

# The State Of Set-Top Box Viewing Data as of December 2009

---

SUBMITTED BY:

Tim Brooks  
tim@timbrooks.net  
www.timbrooks.net

Stu Gray  
stugrayresearch@aol.com

Jim Dennison  
jdennison@cablemeasures.com  
www.cablemeasures.com

February 24, 2010

# The State Of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence

## Preface

Media Research, Planning, and Buying keep getting more complex: audiences are more segmented and there is more programming available than at any time in the past. At the same time, the number of overall touchpoints is increasing as well. Welcome to the Future.

The Council for Research Excellence (CRE) was founded in 2005 as a group of senior researchers across the industry whose mission has been to “advance the knowledge and practice of methodological research on audience measurement...”

One new technology of great promise today lies in the area of Set Top Boxes or STBs. In the following pages you will find a review that really is the first of its kind to provide such an evaluation. You will read about the basic structure of the STB field: the stages of data processing, the companies involved in each stage, and how they relate to each other. As the study’s authors put it: “To our knowledge, this has never been published in one place, and the resulting Data Hierarchy Flowchart and Vendor Hierarchy Summary are key learnings.”

Although some of the material in this volume is technical, those who compiled the findings of the study point out that STBs are at present really the “Wild West” of the research industry, as new players jockey for position in this fast-changing space. Therefore, I am sure you will find the report to be compelling reading.

We thank Nielsen for the funding to conduct this research, Pat Liguori and the Set Top Box committee of the CRE for driving the project from inception to current report and the authors for their cogent analysis.

Last, but certainly not least, we also thank those STB stakeholders who were willing to participate in sharing their information on a confidential basis. Without your cooperation, this document would not have been possible.

*Mike Hess*

*Carat*

*Council for Research Excellence, Chair*

---

Hardly a day now passes that the field of set-top box (STB) measurement is not mentioned in the trade press, although that was not the case in April’08 when the idea for this study was conceived by a committee of the Council for Research Excellence (CRE). Then, as now, there was a need for information about STB data.

Simply stated, our goal was to collect elemental information about the current state of set-top box measurement that anyone involved in audience measurement would want to know. This learning could be used to inform others about this emerging field of measurement. By broadening the base of those with at least a working knowledge of STB measurement, we would establish a foundation upon which others may build, and from which the potential for greater acceptance, faster adoption, and the establishment of “best practices” might spring.

Participating on the STB Committee are CRE members Michele Buslik (TargetCast), Alex Corteselli (Cox Reps), Susan Cuccinello (TVB), Nancy Gallagher (NBC Universal), Colleen Fahey-Rush (MTVN), Lyle Schwartz (Group M), Noreen Simmons (Unilever) and Ira Sussman (CAB). Also participating are Jon Cogan (OMD), Bruce Tyroler (Scripps Networks), and Susan Nathan (Turner Broadcasting). Richard Zackon serves as the committee’s facilitator.

To conduct the study, three independent consultants -Tim Brooks, Stu Gray, and Jim Dennison – were selected to work as a team, based on the complementary nature of their respective industry experience.

Through the collective efforts of the consultants, the STB committee members, several CRE members and CRE Chair Mike Hess, comes “The State Of Set-Top Box Viewing Data as of December 2009.” My thanks to all involved for persevering through the difficult and time-consuming processes of soliciting and collecting data from our target companies. We and the industry owe a debt of gratitude to the fifteen companies that agreed to participate in this study, for it is through the information they provided that knowledge, acceptance, and adoption of STB measurement will grow.

*Pat Liguori*

*ABC Owned TV Stations*

*Council for Research Excellence, Chair, STB Committee*

**The State Of Set-Top Box Viewing Data  
as of December 2009**

STB Committee of the Council for Research Excellence

**Table of Contents**

Preface.....	2
Executive Summary .....	4
Highlights.....	5
Discussion.....	6
Introduction to the Study .....	10
Detailed Findings and Analysis for Aggregators.....	18
Summary for Data Creators .....	26
Summary for Data Owners .....	28
Summary for Third-party Processors.....	29
Appendices Part 1 .....	30
Glossary of Terms.....	31
Data Aggregators Data Table.....	32
Appendices Part 2 (continued in separate document).....	44
Initial Letter to Companies Invited to Participate.....	44
Data Owners Questionnaire .....	44
Data Creators Questionnaire .....	44
Data Aggregators Questionnaire.....	44
Third-party Processor Questionnaire .....	44

# The State Of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence

## Executive Summary

The goal of this study was to conduct a comprehensive examination of how various vendors are using Set-Top Boxes (STBs) to capture and report tuning, and how that data is being processed for use by buyers and sellers. It was very much a learning exercise in a field where little consistent information has been available. The study and this report are not intended to evaluate, or even to reveal, the procedures of any individual company. All participants were guaranteed confidentiality for their individual responses, and findings are shown only in aggregated form.

The study was commissioned by the Set Top Box Committee of the Council for Research Excellence (CRE) in late 2008 and has taken longer to complete than anticipated, for a number of reasons related to this being a new and little understood area for many in the research community. First, we researched the basic structure of the STB field: the stages of data processing, the companies involved in each stage, and how they relate to each other. To our knowledge, this has never been published in one place, and the resulting Data Hierarchy Flowchart and Vendor Hierarchy Summary are key learnings. The five stages of STB data acquisition and processing were defined based on the type of control companies have over the data and the types of activities they conduct. We selected four of the five stages for the study, those categorized into data owners, data creators, data aggregators, and third-party processors. Device makers were not included because data owners were thought to control STB features. Because each stage involves diverse functions, different and somewhat technical questionnaires were designed for each. In total, thirty companies were identified as potential respondents.

Most time-consuming was securing cooperation from companies in this emerging arena where there is much suspicion of motives, secrecy about procedures, and a lack of standard practices. Companies are jockeying for position, making this the "wild west" of research. In addition, because the CRE is funded by the Nielsen Company, there was reluctance on the part of many to provide information. Through the efforts of the consultants, the STB Committee Chair, and the CRE Chair, extensive verbal and written assurances were given that Nielsen funding does not mean Nielsen control of information or direction of the Council. Certainly in this study, Nielsen did not view or receive any of the raw data, and had no influence on any of the questionnaires, data collection, analysis, or report contents.

At the end of an extensive recruitment and placement process, approximately half of the companies approached agreed to participate for the benefit of the industry. Judged against the skepticism expressed by some about the study at its inception, and that this study is the first of its kind, it represents a step forward in our understanding of STB measurement. The companies that participated deserve the gratitude of everyone who is seeking to better understand STB data.

Cooperation was obtained from some companies within each stage. Aggregators responded in sufficient numbers to allow us to report detailed aggregated results without compromising confidentiality. Although all third-party processors responded, none is currently processing STB data, and not enough owners and creators responded to report detailed results, so no data is given for these three categories.

# The State Of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence

## Executive Summary

### Highlights

#### New Learnings

1. Despite fears to the contrary, virtually every type of data that could be obtained from a set-top box *is* currently being obtained and processed by at least one aggregator. There are no significant gaps in what is possible with current STB technology. Data now being processed includes:
  - a. Channel change
  - b. Muting
  - c. Program Guide use
  - d. VOD
  - e. Picture-in-Picture
  - f. Polling
  - g. DVR playback and trick mode (if DVR is internal to the STB)
  - h. Where there are multiple channels on the screen simultaneously (a "mosaic" channel), which one has active audio.
  - i. There are even initiatives underway to address the issue of viewer (or at least household) demographics.
2. In terms of processing procedures, there are creators and aggregators that:
  - a. Record activity at one or two second intervals
  - b. Upload data continuously
  - c. Keep track of channel and program lineups
  - d. Identify commercial breaks
  - e. Can identify reporting STBs down to the household level
  - f. Have implemented some type of quality control
3. At the same time, no single aggregator is doing everything listed above. There is virtually no uniformity between aggregators in terms of data obtained or processing rules. Everyone does it differently. There are no standards.
4. Because of competitive pressures and suspicion of motives, there is a considerable reluctance by many companies to reveal detailed information on their procedures. However this does differ by category. Third-party processors and aggregators (with some notable exceptions) are the most likely to share such information; data creators less likely; and data owners (MSOs, satellite operators, etc.) the least likely to do so.
5. The first reaction from the majority of companies approached was not to cooperate. No one wanted to "go first." However over time, and with persuasion and assurances of confidentiality, many (especially aggregators) came on board. The lesson, we believe, is that it is possible to bring greater transparency to the area of STB procedures. What will be required is a sustained effort, market pressure, and offering clear incentives ("what's in it for me?").

# The State Of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence

## Executive Summary

### Discussion

#### Data Aggregators

The results for aggregators are rich and quite interesting. In many ways, this is the most important data set, as these are the companies that prepare STB data for end use by buyers and sellers. Bearing in mind that not all aggregators responded, and that this is, therefore, a sampling not a census of the industry, two overarching findings emerged:

- There is currently little uniformity in edit rules and reporting conventions; everyone does it a little differently.
- Nearly all types of data internal to the STB are currently being processed by at least some aggregators, and many types by all or almost all aggregators. There seem to be few glaring "holes" in terms of what aggregators are currently able to do, giving us cause for optimism regarding the establishment of "best practices" as STB measurement evolves. Even the concept of viewer demographics is being addressed by some.

Among the specific findings were:

- Channel changes are processed by all.
- Half or more of the responding aggregators currently receive and process muting, program guide use, and VOD, while only a few receive picture-in-picture or polling activity.
- Half process DVR playback when the DVR is internal to the STB, including trick mode (fast forward, pause, and rewind).
- About the only major type of video activity not received by any of the responding aggregators is the use of externally connected devices such as external DVRs, DVDs, VCRs, or game consoles. This is not surprising since software in the STB is the means by which measurement occurs.
- Data is generally received at one or two second intervals and this threshold can be changed either within the reporting system or by reprocessing.
- Frequency of HH data uploads varies widely but is generally daily or more frequently. Most said this frequency could be changed if necessary.
- Aggregators are aware of the need for tuning measurement at the commercial-level and most say they can provide it, in some form.
- Although most aggregators keep track of program and channel lineups, the rigor (e.g., accounting for pre-emptions) seems to vary.
- The smallest unit of geography that can be reported varies widely, from broad DMA all the way down to the individual household.
- Information about households, such as whether or not they are seasonal, and the location of the STB within the household, is generally not known, and therefore not reported. Also unknown is whether a location is residential or commercial.

# The State Of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence

## Executive Summary

- Most aggregators report the demographic characteristics of STB households but not of residents or viewers. However, there are several initiatives underway to address this issue, though opinions vary about whether it needs to be addressed. Some argue that tuning can be associated with product usage or other measures of advertising effectiveness, rendering demographics of less importance.
- Indications are that quality assurance processes are in use by most respondents.
- Although the use of weighting to project to broader universes is limited, some aggregators indicated that this is being considered for the future.

## Third-party Processors

Third-party processors were willing to cooperate, but the results for that group can be simply stated: none are currently processing STB data.

## Data Creators

The results for creators –companies that provide the data-gathering software– can only generally be stated due to limited cooperation from these companies.

Among those that did respond, there seems to be greater consistency than there is among aggregators.

- All or most process channel change, muting, program guide use, DVR playback and trick mode, games, polling, interactive advertising, picture-in-picture, mosaic channel sound, and remote control button presses.
- Data uploads occur at least daily, data is captured to the second and received and processed at that level.
- Quality control procedures are in place.

It is difficult to know if this is representative of all data creator software because of the number of creators who chose not to participate in the survey. If it is, that would suggest that variability in procedures is more likely to occur at the aggregator level than at the raw data-gathering stage. Additional research into this question is needed.

## Data Owners

Very little information can be provided for data owners--companies that deal directly with TV subscribers, own the network or control the return path, and in many cases own the equipment. There were too few responses to even generalize about capabilities, or methodology.

**The State Of Set-Top Box Viewing Data  
as of December 2009**

STB Committee of the Council for Research Excellence

**Executive Summary**

**Conclusion**

All reported findings must be considered preliminary due to the lack of cooperation by some companies in the field. In addition, those who participated were given the option not to answer specific questions that they felt were proprietary. Most, however, were quite forthcoming.

It is hoped this study will engender greater cooperation in the future by demonstrating the value of a non-partisan look at the STB arena, and that such a study can be conducted without harming the interests of any individual company.

**What We Still Need to Learn**

1. More detail is needed on initiatives to measure the demographics of households and viewers. We asked about this generally in aggregator questions 13-15 (e.g., "please list [your] methods"), but most gave a general answer or no answer. A deeper dive into this area is needed. It would be valuable to explore not only what aggregators are doing, but what they plan to do in the future regarding demographics.
2. Although we approached major players in each of the STB areas, valuable learning could doubtless be obtained from non-mainstream media research companies and from smaller operators (e.g. Sunflower) that work in the STB space.

# The State Of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence

## Executive Summary

### Possible Business Implications

1. As long as users are willing to close deals with companies that refuse to reveal their methods in any meaningful way, there will be no guarantee that STB data is valid or actionable. The veil of secrecy surrounding STB data processing will only be removed through buyer pressure. *The industry should insist on disclosure and third-party verification.* Third-party verification, for example through the Media Rating Council, should be a prerequisite to buying this or any other currency-related data, or relying on it in for targeted advertising.
2. Users should consider carefully what kind of data they really want from STBs, since providers ultimately will cater to the needs of their customers. *We should think about what data is relevant.* Are traditional viewer demographics really necessary, or is tuning combined with some measure of product awareness, message reception, or product purchase more relevant?
3. *Prepare to handle a lot more data.* Our survey showed that second-by-second data is being collected 24/7 by several sources, while trade articles cite the very large samples of tens of thousands to millions. At some point this data will enable viewing analyses that aren't practical today, but there could be significant impact on staffing levels required to handle the data and to analyze which measures are meaningful.
4. Related to no. 3, aggregators need to develop user interface systems that can put the huge amount of data being generated into actionable form. Too often this is an afterthought. *Focus on delivery systems.* Given how much data we found is being generated, planning for this cannot begin too soon.

# The State Of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence

## Introduction to the Study

### Goals

The Set Top Box Committee was established by the Council for Research Excellence in April, 2008. Its mission, as stated on the CRE website, is "to conduct a comprehensive examination of how various vendors' STBs capture and report tuning, from which will emerge learning as to the viability of using STB data as a measure of video tuning behavior" ([www.researchexcellence.com](http://www.researchexcellence.com)).

It was not a goal of the study to produce detailed comparisons of companies, or a detailed description of the various aspects of methodology being used for STB measurement or to rate one methodology as superior to another. The primary focus was on current activities and capabilities, with only a few questions about future plans.

### ***Overview of How STB Data is Gathered and Reported.***

Our first task, once we were engaged in late 2008, was to research publicly available information and hold private conversations with knowledgeable sources in order to prepare an overview of how STB tuning data makes its way from individual viewing homes to advertising buyers and sellers and other business users. We also needed to identify the specific companies that were involved in each stage of this data handling, and what exact role each company played. To the best of our knowledge, such a "roadmap" to STB data gathering and handling had never been assembled. We developed two documents for the Committee: a Data Hierarchy Flowchart, showing the stages of data handling, and a Vendor Hierarchy Summary, showing where various vendors fit into the data handling stages.

Figure 1: Data Hierarchy Flowchart (see page 12) graphically illustrates how STB data passes through five stages before reaching its end users. Companies at each stage determine the nature and detail of the data that is available to the later stages, and therefore to eventual users.

The first stage consists of **Data Owners** — cable MSOs, satellite providers, telcos, and a few others — who control the path to viewers' homes and the ability to obtain STB data from them. Because they have a direct relationship with consumers, these entities "own" the data that is gathered.

Stage two is **Device Makers**, the STB manufacturers and the middleware providers that provide the operating software for the boxes from which data is collected. STBs aren't sold directly to consumers, however. Most are provided on a rental basis by the cable, satellite, and telco operators, who decide which devices and which features will be available. An exception is the TiVo DVR, which is sold directly to consumers.

Stage three, which we called **Data Creators**, consists of companies that deploy special software that collects STB level tuning and other clickstream data from the data owners. This software is separate and distinct from the operating software and middleware that runs the boxes, and generally sits on top of those applications.

# The State Of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence

Stage four is **Data Aggregators**, companies that receive the raw information from the data creator applications, apply edit rules, and otherwise process it for end users. These are basically research companies that are active in the STB field.

Stage five, **Third-party Processors**, consists of two types of companies, both of which are engaged by end-users and through which they get access to the output from data aggregators. *Software System Providers* are companies that edit, combine, and format data to the exact specifications of the end-user. *Database Vendors* are companies that have databases segmented by various characteristics that can be correlated to STB data through either ZIP code or other data points, thus expanding the characteristics by which STB data can be analyzed. Database vendors can be involved in the process at stages prior to this one, and can work with data owners, creators, and aggregators.

In our discussions during the design of the study, we determined that the second stage, **Device Makers**, while integral to creating set-top box devices with software capabilities that enable STB measurement, was controlled by requirements given by the data owners and the functionality enabled by the data creators. Given this, and that our primary focus in the study was on current capabilities, we decided not to send questionnaires to the device makers. We felt that the capabilities of currently deployed STB devices would be reflected in the questionnaires received from data owners, data creators, and data aggregators.

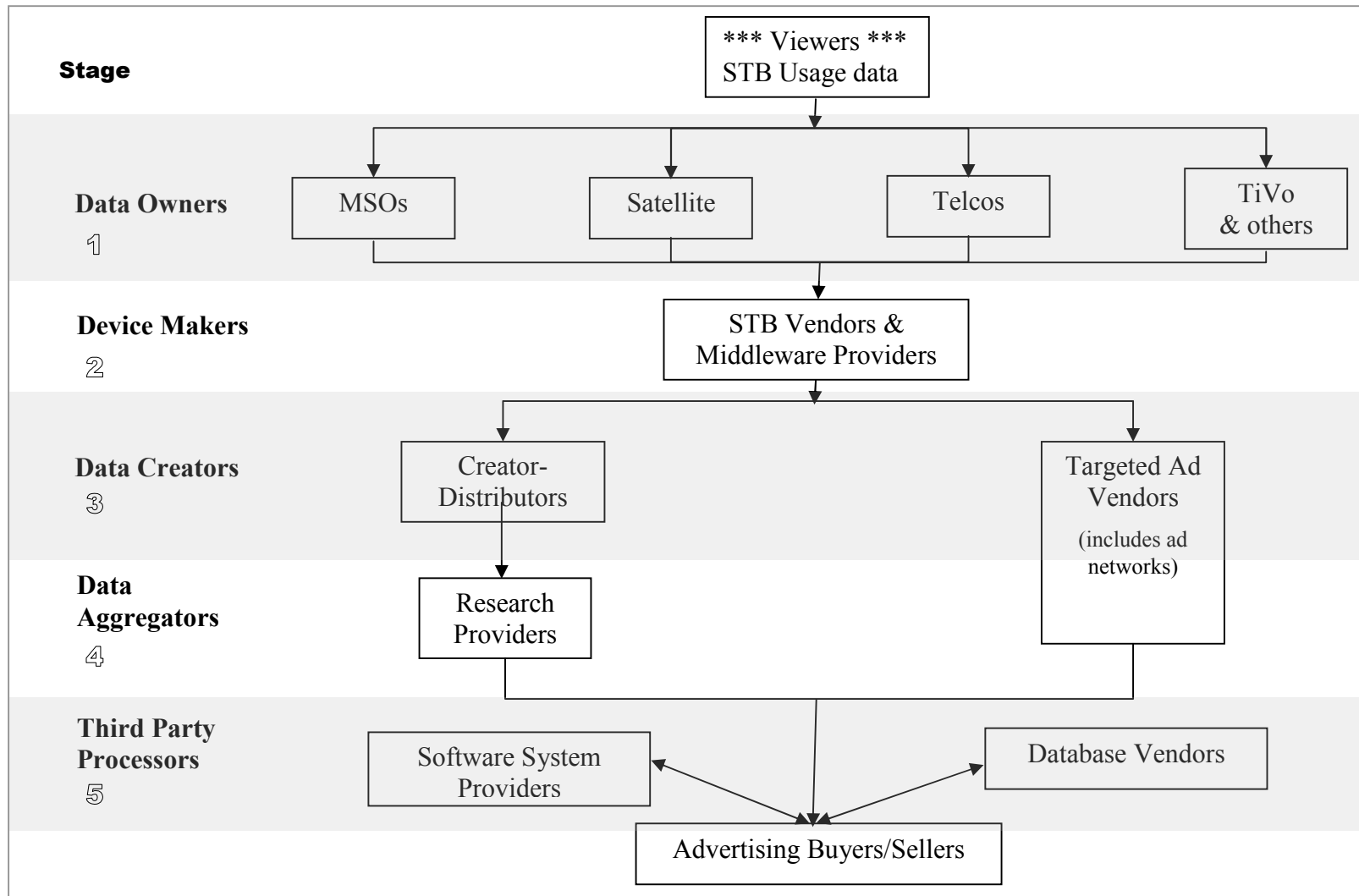
Figure 1: Data Hierarchy Flowchart, page 12, shows the relationship of the five stages and the categories of companies in each.

# The State Of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence



**Figure 1: Data Hierarchy Flowchart**



# The State Of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence

Within these categories, we worked with the STB Committee to create a list of the companies that have marketplace presence, rather than a complete list of all companies that are active in each space. Table 2: Vendor Hierarchy Summary, page 14, shows many of the companies that are active at each of these stages, with some active at multiple stages. TiVo, for example, is a data owner (has direct access to its STB boxes), device maker (sells DVRs), data creator (creates its own software), and data aggregator (processes and sells the data). Others operate at a single stage; for example, TNS, Rentrak, and Nielsen are data aggregators who obtain STB data from data creators, with the permission of the data owners.

Targeted ad vendors such as Invidi, Visible World, and perhaps Canoe Ventures, are interesting cases as they could be both creators and aggregators, since they process data only for their own use for ad placement and proof of performance, and *may* or may not sell it to outside entities.

With the concurrence of the STB Committee, we excluded device makers from the study, for the reasons given above (see page 11), and database vendors (e.g. Axciom, Experian, Claritas), since they do not deal directly with STB usage data although their data could at some point be fused with it. Both could be canvassed at a future date.

In all we identified 30 companies that we and the Committee believed were active in one or more stages of STB data processing. Of these, 15 agreed to provide information, although five turned out not to be currently creating or processing STB data.

We would like to publicly thank the following for their willingness to provide information for this fact-finding study, done for the benefit of the industry:

- Canoe Ventures
- Concurrent
- Cox Media
- Donovan Data Systems
- Google
- IMS
- The Nielsen Company
- OpenTV
- Rovi
- Star Media
- Telmar
- TiVo
- TRA Global
- TVWorks
- Visible World

(some completed more than one questionnaire)

**Table 1: Companies Invited to Participate**

AT&T U-verse	Donovan Data Systems	Rovi
Bright House Networks	Google	Star Media Enterprises
Cablevision	Guideworks	Telmar
Canoe Ventures	Invidi	Time Warner Cable
Charter	IMS	TiVo
Comcast	itaas	TNS
Concurrent	Navic	TRA
Cox Media	The Nielsen Company	TVWorks
DirecTV	OpenTV	Verizon FiOS
Dish Network/Echostar	Rentrak	Visible World

# The State of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence



**Table 2: Vendor Hierarchy Summary**

<b>Category</b> (and examples. Not a comprehensive list)	<b>Definition</b>
<b>1. Data Owners</b> <ul style="list-style-type: none"> <li>MSOs (TWC, Comcast, Cox, Charter, Cablevision, Bright House, others)</li> <li>Telcos (Verizon FiOS, AT&amp;T U-verse)</li> <li>Satellite (DirecTV, Dish Network/EchoStar)</li> <li>TiVo (perhaps others, like Roku )</li> </ul>	Companies that deal directly with TV subscribers, own the equipment and the network, or control the return path (i.e. telephone, cable, fiber)
<b>2. Device Makers – hardware and O/S</b> <ul style="list-style-type: none"> <li>STB mfrs (Motorola, Cisco, Pace, Samsung, Panasonic, Moxi/Digeo, ADB, Echostar Technologies, and other Satellite STB manufacturers)</li> <li>Middleware Providers (Microsoft, OpenTV, TVWorks, others)</li> </ul>	Provide STB hardware and its operating software Make STBs for network operators and direct-to-consumers (TiVo) Software that sits between OS and applications. It extends basic STB functions, manages other software and enables many STB measurement functions
<b>3. Data Creators</b> <ul style="list-style-type: none"> <li>Creator/Distributors (Navic, Rovi, itaas, TiVo, DirecTV, GuideWorks, OpenTV, EchoStar Technologies/Dish, Verizon-FiOS)</li> <li>Targeted Ad Vendors (Invidi (addressable ads), Visible World (addressable ads), Navic (targeted ads, ad network))</li> </ul>	Deploy software that collects tuning data and other clickstream data at the STB level Generate and collect raw data and will distribute data to other aggregators, with Data Owner permission Collect and use tuning data for their own ad placement services. (Also aggregate data) Perhaps won't provide the data to outside research entities
<b>4. Data Aggregators</b> <ul style="list-style-type: none"> <li>Research Providers (TNS, Rentrak, TRA, Nielsen, Concurrent)</li> <li>Targeted Ad Vendors (Invidi, Visible World, Navic, Google (ad planning, placement, and posts), Canoe Ventures (addressable ads, ad network, ad posting))</li> </ul>	Process raw data collected by data creators, package and sell research products and services Databases and reporting/analytical systems for use by buyers and sellers These vendors aggregate and process the tuning data for use in their own services, in some cases in addition to creating data with their own applications.
<b>5. Third-party Processors</b> <ul style="list-style-type: none"> <li>Software System Providers (Donovan, Star Media Enterprises, Telmar, IMS,)</li> <li>Database Vendors (Acxiom, Experian, Claritas)</li> </ul>	Provide research applications using data provided by clients of data aggregators Use processed data and respondent-level data to provide analytical applications and databases. Use STB subscriber data to characterize subscribers and fuse STB data with other data, like home ownership, income, etc., or to assign target profiles.

# The State of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence

## ***Methodology of the Study***

Since the types of activities are different at each stage, we developed four questionnaires, tailored to owners, creators, aggregators, and third-party processors (see appendix Part 2). Admittedly, there was some overlap, but the functions within each stage were sufficiently different that using a single questionnaire would have been confusing. The questionnaires were edited and finalized with input from STB committee members. We believe the questionnaires can serve as useful templates for future evaluation and analysis of providers in the STB field.

It took considerable time and effort to identify and contact the key people to approach within each company, and sometimes required the help of committee members. In most cases, the first approach was an overnight letter from the CRE explaining the purpose of the study, introducing the three consultants, and requesting cooperation. The consultants then followed up by sending the appropriate questionnaires by email.

An important element of the methodology was the guarantee of respondent data confidentiality. This was absolutely essential because, in conversations with the consultants, executives in the field had made it clear that they were extremely concerned about competitors gaining an advantage by learning the details of their capabilities and confidential processing rules. All companies approached were provided with the following written assurance:

**"The Council for Research Excellence wants you to know that neither the CRE members, nor the STB Committee members, nor The Nielsen Company employees or contractors, will have access to the responses to this questionnaire. Access to the completed questionnaires will be given only to three consultants to the Council (Gray, Brooks, and Dennison) for the sole purpose of analyzing the responses and reporting the responses to the Council in an aggregated form. The consultants will destroy the questionnaires within three months of delivering the report."**

In addition, respondents were given the option to not answer questions they felt would require them to reveal proprietary information, or to answer certain questions at a later date. Most respondents were fairly forthcoming. Every respondent answered at least some questions and most answered nearly all questions.

The guarantee of confidentiality will be rigidly adhered to. All findings contained in this report will be presented in aggregated form only, and we will not reveal responses from any individual company without its explicit permission. The purpose of this investigation is to shed light on the practices being followed in the STB field as a whole, not to reveal or pass judgment on the procedures of any individual company.

# The State of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence

## **Barriers Encountered; How Greater Cooperation Might Be Secured in the Future**

Despite the aforementioned guarantee of confidentiality, and the offer of consultant NDA's if desired, the level of suspicion and reluctance to share information was quite high. Some companies, understanding that the study was non-partisan and for the good of the industry, readily cooperated; others required considerable persuasion and still others refused to cooperate. The reasons varied by company and by category, but generally:

- Some felt there was "nothing in it for them."
- Some said their procedures were still in development, and they did not want to talk about them "at this time."
- Quite a few were suspicious of Nielsen's connection with the CRE, and worried that, despite the guarantees of confidentiality, Nielsen would somehow gain access to their specific responses. This is emphatically not the case.
- In a number of cases, some executives within the company saw the value of such a study and were anxious to cooperate, but were blocked by management, legal or other departments.
- Some said there were "Too many committees asking for [similar] information", referring to groups such as CIMM, Collaborative Alliance, CRE, Starcom Digital Group, etc., and felt the various groups should get together.
- Concern was expressed that publicity about STB data might raise questions with the public about viewer privacy and anonymity.

In the end, completed questionnaires were received from all four categories (owners, creators, aggregators, and third-party processors). Aggregators responded in sufficient numbers to assure anonymity when combined together, and that category is reported in detail in the following pages. All of the third-party processors that were contacted were willing to cooperate but none are currently processing STB data. The same situation applied to one data creator. An insufficient number of owners and creators responded to assure anonymity, therefore limited information about the state of STB processing in those sectors is provided in the summaries that begin on page 26.

In terms of securing greater cooperation in the future, we hope this report will demonstrate the value of a non-partisan look at procedures in this emerging field, and that such a study can be carried out without harming the interests of any company. In fact, it hopefully will bring value to providers in the field by making buyers and sellers more comfortable with the data they are being asked to trust. Ultimately, only market pressure will compel research companies that prefer to remain secretive to share their procedures with broad-based buyer/seller groups such as the CRE, and ultimately, to seek accreditation from the Media Rating Council.

**The State of Set-Top Box Viewing Data  
as of December 2009**

STB Committee of the Council for Research Excellence

***Timeline of Events***

- August 2008 - RFP released.
- November 2008 - Project awarded to consultants, Brooks, Gray, & Dennison.
- December 2008 - Strategy meeting with STB committee; methodology finalized.
- January-March 2009 - Researched the field, developed and refined questionnaires, identified contacts.
- April-May 2009 - Contacted key executives at 30 companies; sent them questionnaires.
- June-November 2009 - Follow-up and meetings with potential respondents.
- December 2, 2009 - Study closed.
- December 2, 2009 - Status report by STB Committee Chair, Pat Liguori, at a meeting of the Collaborative Alliance (NYC).
- December 21, 2009 - Draft report submitted to Committee for comments.
- February 5, 2010 - Final report delivered to Committee.
- February 24, 2010 - Review of final report completed.

# The State Of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence

## Detailed Findings and Analysis for Aggregators

### *General Placement and Completion Status*

Data aggregators process raw data collected by data creators, producing standard measures such as reach, frequency, ratings, and total audience in addition to data and measures they create with proprietary applications. Aggregators may provide the standard measures to buyers and sellers as unformatted data sets, or package and sell access to the data through research applications and services, such as databases and reporting/analytical systems. Some vendors in this category have ad networks, or targeted addressable ad placement products so they aggregate and process the tuning data for use in their own services.

We identified eleven companies in this category based on their believed activity in STB measurement, their market position, and industry awareness. The companies, listed alphabetically:

- Concurrent (formerly Everstream)
- Canoe Ventures
- Google
- Invidi
- The Nielsen Company
- Rentrak
- Rovi (formerly Macrovision; also a data creator)
- TiVo (also a data owner and creator)
- TNS
- TRA Global
- Visible World

### *Completion Status Summary*

- Canoe Ventures was not processing STB data, leaving ten in this group.
- 7 of 10 completed data aggregator questionnaires.

# The State of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence

## **General Observations for Types of Data Collected and Use of Supporting Databases**

**(based on Data Aggregator Questions 1 - 8)**

*The Data Aggregator questionnaire is incorporated into a data table that begins on page 32.  
The question numbers referred to here correspond to that table.*

- We found little uniformity in edit rules and reporting conventions. This is perhaps because of the lack of industry standards and the secrecy of aggregators regarding their procedures.
- Nearly every type of data we asked about is being processed by at least some aggregators, however some is processed by most, and almost never is it processed by all.
- The pattern of non-responses also suggests a lack of universally accepted standards and knowledge of what others are doing. Most respondents did answer most questions. However, for each specific question, one or two respondents declined to answer, suggesting lingering paranoia, concealment of what is perceived to be proprietary methods, or that they have yet to satisfactorily address an issue.
- Nevertheless, for several reasons there is cause for optimism regarding the establishment of “best practices” :
  - 1) aggregators currently are able to receive and process most types of data internal to the STB;
  - 2) the data is collected from households daily (or more frequently);
  - 3) tuning is captured to the second, and delivered to aggregators at that granularity, and,
  - 4) aggregators are making serious efforts to keep track of program and commercial content and channel lineups.

### **1. What types of STB activity are aggregators capturing and processing into reports?** *(refer to Questions 1&2)*

- All responding aggregators currently receive and process channel changes. This is the only STB activity that is universally received and processed.
- VOD, muting, and program guide use were received and processed by a slight majority of the aggregators who responded. Other activities, such as trick mode or Picture-In-Picture, were received and processed by half or less.
- Most (4 of 7) did not answer the question about audio on mosaic channels, suggesting that it may not be widely captured. Mosaic channels are those with multiple channels shown in small windows on a single screen, often for the purpose of selecting channels.
- If data about an activity is received, it is almost always processed. The only exception was mosaic audio, which was not necessarily processed by those who received it.

## The State of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence

- Perhaps the most interesting and important finding overall is that while many of these activities currently are not being received and processed, most of them can be, as evidenced by the fact that at least some aggregators are doing so. The major exception is externally connected devices (DVR, game consoles, etc.), which currently appear to be “off the radar.”

### 2. How often is raw data updated and uploaded? *(refer to Question 3)*

- Frequency varied widely, from “continuously” to “several times a day,” “daily” and “once per week per STB.”
- The most frequent answer was at least daily (5 of 6).
- Nearly all (6 of 7) said that the frequency of update/upload could be changed.
- Bottom line, there does not appear to be standardization here but aggregators said they can adjust this to meet clients’ needs. The only barrier cited was cost.

**Table 3: Summary of Received and Processed Data (Questions 1-8)**

Activity	Received/Processed by:
Channel change	All (6 of 6)
VOD	Majority (3 of 5)
Audio on mosaic channel	Majority (2 of 3 who answered)
Trick mode (fast forward, pause, rewind)	Half (3 of 6)
Muting	Majority (3 of 5)
Program Guide Usage	Majority (3 of 5)
DVR playback with program titles (where DVR is built in to the STB)	Half (3 of 6)
Picture-in-Picture (channel, main picture, audio)	Very few (1 of 5)
Polling	Minority (2 of 5)
Games	None
<b>Externally connected devices:</b>	
DVR	None
DVD	None
VCR	None
Game console	None
Internet video via STB e.g. Netflix, Roku, SlingCatcher, Apple TV, Vudu	None

*Note that some questions had 1 or more respondents who declined to answer.*

### 3. What is the shortest duration tuning entry that can be reported? *(refer to Question 4)*

- Most said one or two seconds (4 of 6). One indicated that it was event-based, with “no limit generally.”
- Aggregators were split on whether this threshold could be changed within the reporting system. However, of those who said “no” (3 of 6), all indicated that it could be changed by reprocessing the data.

## The State of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence

- Our conclusion here is that aggregators are processing very short tuning durations, and can adjust that duration if needed.
- 4. Do aggregators maintain their own databases of program lineups?** *(refer to Question 5)*
- A majority (5 of 6) maintain a lineup database. Sources include internal records, syndicated sources, Tribune Media Services, and TNS.
  - The one that does not maintain a database uses the program codes supplied in the raw data records.
  - Most (3 of 4) update their lineup information for pre-emptions and overruns.
  - All adjust for different time zones, where necessary.
  - The summary here is that most aggregators are keeping track of lineups, although the rigor (e.g. accounting for pre-emptions) seems to vary.
- 5. Can aggregators identify tuning to commercial content?** *(refer to Questions 6-7)*
- A majority (5 of 7) said they could identify commercial content; the remainder could not.
  - Of the five who could identify commercials, all said they could do so to the second. A majority (4 of 5), with one not answering, said they could differentiate national commercial content from local.
  - In summary, aggregators are aware of the demand for commercial-level data and most say they can provide it, in some form. Since this has long been one of the basic selling points of STB data, it is clearly a goal for more to do so.
- 6. Are aggregators keeping track of channel lineups, and if so, how?** *(refer to Question 8)*
- A majority (4) said they did keep track of channel lineups, although two declined to answer this question.
  - Those who did used internal records, syndicated sources, and TNS.
  - Those who did also claimed that channel lineup changes were reflected “immediately” or, at the least, daily.
  - Most (3 of 4), however, said that they were *not* able to determine network/channel carriage down to the individual STB level. Therefore, they could not calculate coverage area ratings.
  - In summary, channel lineup information is basic to STB data reporting, and aggregators appear to recognize this and be dealing with it. Some, however, were reluctant to answer this question, suggesting that it is still an issue for them.

### **General Observations on Use of Geographic and Demographic Characteristics of Households and Viewers**

*(based on Data Aggregator Questions 9 – 15)*

- 1. What is the smallest unit of geography reported?** *(refer to Question 9)*
- There does not appear to be uniformity with regard to the smallest unit of geography that can be reported. Currently, the minimum seems to range broadly from DMA all the way down to Household.

**The State of Set-Top Box Viewing Data  
as of December 2009**

STB Committee of the Council for Research Excellence

**2. Are Non-residential and seasonal households differentiated?** *(refer to Questions 10-11)*

- Residential and non-residential data are combined by 4 of 6 who answered. Only one respondent identifies non-residential data in such a way that users of reports data can differentiate it from residential subscribers' data.
- Seasonal households are not being identified by any of the three who answered.

**3. Is the room location of the STB known/reported?** *(refer to Question 12)*

- Identifying sets within the household by their room location is not happening yet.

**Table 4: Summary of STB Characteristics and Demographics Usage** *(Questions 9-15)*

Question/Issue	Responses	Low Response
Smallest reportable geographic level	Big range- from DMA down to HH	
Differentiation of non-residential HH	A majority (4 out of 6) combine residential and non-residential households	
Identification of seasonal households	None (0 out of 4) have identified seasonal households	R
STB Location within the household	None (0 out of 6) have identified location within the household	
Household demo characteristics	Most (5 out of 6) report HH characteristics; 2 from ascribed data, 3 from other sources	
Demographics of residents	Half (3 out of 6) report demos of residents, one licensing from their data source	
Demographics of viewers	Only 1 out of 5 does; but did not describe how.	

*Note that some questions had 1 or more respondents who declined to answer. A flag denotes where the number of non-responders was 3 or more*

**4. What types of demographics for households and viewers are reported?**

*(refer to Questions 13-15)*

- Most (5 of 6) aggregators are reporting demographic characteristics of STB *households*, such as HH size, presence of children, etc. One of five uses data directly from the source; the rest use modeled or ascribed data. No one indicated that a change of methods is planned for the future.
- Demographics of *residents* within the households are being reported by half (3 of 6), and would be an important metric in modeling viewer demographics.
- Demographics of *viewers* are reported by one; not planned by most (4 of 5).

We did not specifically ask aggregators their opinion about the consequences of the missing metrics noted above. It is our opinion that some media buyers and sellers may feel this represents a problem, while others may argue that these metrics are unimportant as long as household tuning can be linked to product purchase or product usage data, or other measures of advertising effectiveness.

**The State of Set-Top Box Viewing Data  
as of December 2009**

STB Committee of the Council for Research Excellence

**General Observations on Data Quality Assurance  
and Use of Weighting/Projections**

**(based on Data Aggregator Questions 16-26)**

**1. What is the extent of data quality monitoring and editing?** *(refer to Questions 16-22)*

- There are good indications that quality assurance is being done by most respondents. Most of the aggregators say they control the criteria for quality of data that flows into the reporting and analytical systems.
- Nearly every type of quality control process we asked about is being done at some level by at least some aggregators, although no one is doing every type.
- The responses indicate an awareness of the important issues, and an attempt to address them, if not now, then in the future. However, findings for certain questions are limited because every aggregator declined to answer at least one question, and every question had at least one no answer.

**Table 5: Summary of STB Data Quality Control (Questions 16-26)**

Question/Issue	Responses	Low Response
Percentage of STBs reporting data each day	Answers given were <sup>1</sup> / <sub>7</sub> and 80-85%, and one gave “configurable”. Most (4 of 7) gave no estimate or declined.	R
Criteria for data usability	Most (6 of 7) apply criteria daily or continuously, consistent with the fact that most collect and process data daily. All who answered (6 of 6) said it is they who make the decision on criteria for data usage, rather than another entity.	
Specific editing rules (STB on/set off, Dwell time, Time Zone Adjustment)	Most (6 of 7) declined to answer, or left it blank. One provided some detail on edit rules.	R
STB malfunctions	A few (3 of 7) said they have some process in place to identify malfunctions, and one said they do not have the data they need to resolve some of these issues	
STB loss of data, differentiating no data from non-viewing	Three gave examples of malfunctions that could result in loss of data. Majority did not answer.	R
Keeping records of malfunctions, reporting error levels in reports	<ul style="list-style-type: none"> <li>- Two said they do keep records (2 of the 3 providing answers), and the other does not.</li> <li>- Only one publishes errors in reports</li> <li>- Majority declined to say or left it blank (4 of 7)</li> </ul>	R

*Note that some questions had 1 or more respondents who declined to answer. A flag denotes where the number of non-responders was 3 or more*

## The State of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence

*A note on how the overall completion rate for this section affects the conclusions: Unanswered questions and a lack of detail in answers to several questions possibly affect the usability of this section. Several respondents marked “Decline to Answer” or just left the answer blank for some of the more sensitive and technical questions about methodology. Perhaps this is because of the competitive marketplace, as mentioned previously. For the questions with low response, caution is advised in extending the reported general observations about methodology to the group as a whole.*

### 2. Is weighting being used to correct for under representation and for projecting to a larger population? *(refer to Questions 23-26)*

- The use of weighting is limited, perhaps because data about subscriber characteristics is limited by the privacy requirements for STB data collection. Most data is being collected anonymously, not through a panel where the characteristics of a household, such as race and ethnicity would be known. One aggregator responded that weighting is being done to adjust for household characteristics, such as geographic skews and distribution of STB models.
- The use of universe estimates to project the collected data to some larger population, such as a DMA or to the U.S., or to produce ratings for some population, is being done by half (2 of 4) of respondents at this time, although others indicated that these are future enhancements being considered.

**Table 6: Summary of STB Weighting and Projections** *(Questions 16-26)*

Question/Issue	Responses	Low Response
Weighting the data	One (1 of 5) does, on several variables; most (4 of 5) do not	R
Projecting the data to a larger population or universe	Few (2 of 7) do, but most do not.	R
Updating statistics of the STB population	Few (3 of 7) do it monthly or quarterly, and one also said annually	
Reporting ratings based on universes	One of 7 said they may do this depending on the data source and the intention; the rest either do not report ratings, or didn't answer.	R

*Note that some questions had 1 or more respondents who declined to answer. A flag denotes where the number of non-responders was 3 or more*

## **The State of Set-Top Box Viewing Data as of December 2009**

STB Committee of the Council for Research Excellence

In summary, this part of the study illuminates critical questions about how STB data quality issues are being addressed by various aggregators. Respondents provided few details, however. Each seems to be working out major issues as it gains experience with the data and works with clients of the end-user reporting systems. The number of STBs being measured is greater than that of traditional panel or survey data, so the application of editing rules and use of weighting and projection are being developed and applied differently.

The reluctance of the respondents to provide specific answers is not necessarily an indication that they have no processes in place, but is perhaps an indication of their reluctance to divulge innovative methods in an emerging and competitive environment. Aggregators are also limited in what steps they can take to identify and correct errors, since most rely on the data creators to generate the raw data, who in turn rely on the data owners to run critical parts of the data collection operation. And, since many of the aggregators get raw data from the same sources, they all are dealing with similar limitations.

# The State Of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence

## Summary for Data Creators

### **General Placement and Completion Status**

Data creators design, develop, and deploy software that collects tuning and other clickstream data at the STB level. Creators may have other products and services that aggregate the data, report the results, or deliver ads, but this part of the survey focused on the data generated in the STB.

We identified nine companies in this category based on the deployment of their data creation software or their market position and industry awareness. The companies, listed alphabetically:

- DirectTV
- Dish Networks/EchoStar Technologies, Inc.
- GuideWorks
- Itaas
- Navic
- OpenTV
- Rovi (formerly Macrovision; also a data aggregator)
- TiVo (also a data owner and aggregator)
- TVWorks

### **Completion Status Summary**

- 3 of 9 completed data creator questionnaires.
- 1 of 9 currently doesn't create STB data,  
(TVWorks responded that it does not generate STB tuning information, and thus did not complete a questionnaire.)

### **Details**

- Because we did not receive enough completed questionnaires to report aggregated numbers without the risk of certain responses being linked to one of the respondents, ***we are not reporting data for this category***, but have summarized some findings, below.

### **General Observations on Data Collected and Completeness of Responses**

- It is clear that among the group that responded, there is great consistency in what is being collected. However, the number of non-participants raises the possibility that other data creators may have different data collection capabilities.
- There is uniformity in the way data is being collected and processed, especially the collection of tuning-related events, and DVR activities when the DVR is built into the device being measured.

## The State of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence

- All or most who responded capture channel change, muting, program guide usage, DVR playback, DVR trick mode, games, polling, interactive advertising, picture-in-picture, mosaic channel sound, and remote control button presses. Few are collecting VOD usage. None are collecting external DVD, VCR, or game console usage.
- Those who responded indicated that uploads of STB data from households occur at least daily and that the frequency is configurable.
- Granularity to the second is captured and quality control procedures (error handling, malfunction reporting, and data recovery) are in place.
- Most respondents answered most questions, leaving just a few unanswered. The most sensitive question not answered was about the percentage of STB's that report data on a daily basis, a key quality measure in most respondent-based research.

In summary, this part of the study sheds some light on critical questions about what data is being collected and how STB data quality issues are being addressed, at least by those who responded. Many details are missing, and certainly more completed questionnaires would have been beneficial.

Since most of the data creators rely on the data owners to run critical parts of the data collection operation, data creators, like data aggregators, are limited in what steps they can take to identify errors and correct them. This will be an ongoing challenge as data owners deploy multiple applications to accommodate future growth of tru2way and IPTV operating environments. Some of these newer STB applications will run simultaneously, and many will contain tuning collection capability necessary for the enhanced features built into the application. It will be up to the data owners, data creators, and data aggregators to work out methods for monitoring data quality, and avoiding conflicts between applications that might result in loss of data.

# The State Of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence

## Summary for Data Owners

### ***General Placement and Completion Status***

Data owners control the return path from subscriber households within which set-top boxes collect tuning and other clickstream data. In this case, an STB is broadly defined to also include stand-alone devices like TiVo, which may be connected to an STB.

Although data owners may have other products and services that aggregate data, report results, or deliver ads, and may have close relationships with vendors who are deploying ad networks, or interactive content on the STB, this questionnaire focused on how STB measurement was being deployed and managed.

We identified eleven companies in this category, mostly based on their subscriber size and market segment – cable, telco, satellite, and TiVo. The companies, listed alphabetically:

- AT&T U-verse
- Bright House Networks
- Cablevision
- Charter
- Comcast
- Cox Media
- DirecTV (also a data creator)
- EchoStar/Dish (also a data creator)
- Time Warner Cable
- TiVo (also a data creator and aggregator)
- Verizon-FiOS (also a data creator and aggregator)

### ***Completion Status Summary***

- 2 of 11 completed data owner questionnaires.

### ***Details***

- Because we did not receive enough completed questionnaires to report aggregated numbers without the risk of certain responses being linked to one of the respondents, *we are not reporting detailed data for this category.*

### ***General Observations on the Lack of Responses***

- Based on our discussions with this group during the placement process, we believe all eleven are involved in measurement of STBs. This was supported by what we learned through trade articles and public statements.
- The number of study non-participants is disappointing; it leaves users of STB data in the dark about important methodological and technical issues.

# The State Of Set-Top Box Viewing Data as of December 2009

STB Committee of the Council for Research Excellence

## Summary for Third-party Processors

### *General Placement and Completion Status*

Third-party processors provide software systems, information processing services, and research applications using data provided by the clients of data aggregators. They use processed data and respondent-level data to provide analytical applications and databases to the advertising industry. Also included in this group are database companies that use STB subscriber data to characterize subscribers and fuse STB data with other data, like home ownership, income, etc., or to assign target profiles.

We identified seven companies in this category, mostly based on market presence. Four are software system providers and three are database vendors. Based on the belief that the database vendors were not yet using STB tuning data in their profiling systems, we decided not to contact them for this study.

The companies, listed alphabetically:

#### Software System Providers

- Donovan Data Systems
- IMS
- Star Media Enterprises
- Telmar

#### Database Vendors (Not contacted)

- Acxiom
- Experian
- Claritas

### *Completion Status Summary*

- 4 of 4 were willing to participate, but none are processing STB measurement data at this time.

### *Details*

- Because the companies we contacted are not yet processing STB data, we have no specific data to report for this category.

### *General Observations*

Based on our discussions with this group during the placement process, we learned that because third-party processors currently have no STB data, they are not working on integrating STB measurement data into the reporting and analytical systems that are an integral part of the advertising buying and selling process. They seem eager to obtain and run test data so that their systems won't lag behind when production data becomes available.

**The State Of Set-Top Box Viewing Data  
as of December 2009**

STB Committee of the Council for Research Excellence

**Appendices Part 1**

Glossary of Terms

Data Aggregators Data Table

**Appendices Part 2**

Initial Letter to Companies Invited to Participate

Data Owners Questionnaire

Data Creators Questionnaire

Data Aggregators Questionnaire

Third-party Processor Questionnaire

**The State Of Set-Top Box Viewing Data  
as of December 2009**

STB Committee of the Council for Research Excellence

**Glossary of Terms**

*(selected terms used in this study)*

<b>Data Aggregators</b>	Companies that process raw STB data (collected by data creators) into standard measures, and then package and sell research products and services, e.g., Nielsen, and TRA Global
<b>Data Creators</b>	Companies that deploy software that collects tuning data and other clickstream data at the STB level, e.g., OpenTV, Rovi, and TiVo
<b>Device Makers</b>	Companies that provide STB hardware and its operating software, e.g., Motorola, Cisco, and EchoStar Technologies
<b>Data Owners</b>	Companies that deal directly with TV subscribers, own the equipment and the network, or control the return path (i.e. telephone, cable, or fiber), e.g., Cox Media, Verizon FiOS, and DirecTV.
<b>(IPTV) Internet Protocol Television</b>	A system through which digital television content is delivered using the architecture and networking methods of the Internet., e.g., the Internet and broadband Internet access networks, instead of being delivered via traditional over-the-air broadcast, satellite signal, and cable television formats
<b>Modeling (data)</b>	The process by which data, that is unobtainable due to cost or logistical practical matters, is estimated or predicted using mathematical formulas based on observed or theorized relationships to measurable data
<b>Mosaic</b>	A collection of channels displayed simultaneously as thumbnails, on a single TV screen, usually for the purpose of selecting a channel to watch
<b>Node</b>	A collection of cable homes in a particular area that shares the same coax or fiber network distribution equipment, and in some systems may be the smallest addressable entity.
<b>Set-top box (STB)</b>	A device that connects to a television and an external source of signal. It receives and decodes video and/or audio signals into content which is displayed on the television screen. For this study, standalone DVRs like TiVo, were considered STBs.
<b>Targeted vs. addressability</b>	Addressability is the ability to send content to a particular node, household or STB. Targeting is the ability to segment subscribers by certain criteria, and deliver ads or content to the segments, using some level of addressability.
<b>Third-party Processors</b>	Companies that provide research applications using media usage data provided to them by the clients of data aggregators. Some of these companies may also provide segmentation databases.
<b>tru2way</b>	The brand name for a software platform that permits the integration of set-top box functionality into TVs and other viewing devices. According to the National Cable and Telecommunications Association, tru2way is a Java-based platform that serves as a universal translator for interactive TV guides, video-on-demand, games and other new interactive (two-way) applications
<b>Video on Demand (VOD)</b>	A system which allows users to select from an available library of video content, either for free or for a fee. Television VOD systems either stream content through a set top box, allowing viewing in real time, or download it to a device such as a computer, digital video recorder or portable media player for viewing at any time

**The State Of Set-Top Box Viewing Data  
as of December 2009**

STB Committee of the Council for Research Excellence

**Data Aggregators Data Table**

The following table is laid out in the format of the questionnaire, with the aggregated responses given for each of the questions.

Where respondents provided comments in the questionnaires, they are included, verbatim, unless confidentiality was an issue. In some cases the verbatim comment was annotated to improve readability.

The questionnaire contained two columns, *Answer Later*, and *Decline to Answer* that respondents could use if a response would require proprietary information. In some cases, respondents left a question blank, and did not use the *Answer Later*, or *Decline to Answer* columns. All of these responses are combined in the following table under the column labeled *Declined to Answer, Deferred Answering, or Left Answer Blank*.

Therefore, the sum of the responses shown for a particular question will equal the number of total respondents, which was seven, unless the question is nested, and applies only to a subset, such as those who answered the previous question “yes”.

For question 1 and all of its parts, a blank response in the *Receive* and *Process* columns meant the data type was not received or processed. For this question only, a blank answer was not counted as “no answer”, and the heading reflects that.

## Data Aggregators Data Table

Num.	Please answer the following questions as completely as you can, without giving proprietary or confidential information. For questions you prefer not to answer, please check one of the two columns to the right to tell us whether we should follow-up for a response at a later time.	Yes	No	Comments	Declined to Answer, Deferred Answering,
1.	<i>Which of the following types of STB activity do you receive as raw data and process into reports? (check the appropriate column for each)</i>	<i>Note: A blank response in these columns meant the data type was not received/processed, not a "no answer"</i>			
1a.	Channel Changes	6	6		1
1b.	VOD	3	3		1
1c.	DVR playback (where feature is built into the STB)	3	3		1
1d.	DVR playback with program titles	3	3		1
1e.	Fast Forward	3	3	VOD only <div style="border: 1px solid black; padding: 2px; width: fit-content; margin: 5px auto;">comments refer to one respondent only</div>	1
1f.	Pause	3	3	▪ VOD only	1
1g.	Rewind (Backward Play)	3	3	▪ VOD only	1
1h.	Games	0	0		2
1i.	Short codes (codes entered with the remote control, functioning similar to cell phone CSC codes)	0	0		2
1j.	Polling (Voting)	2	2		2
1k.	Muting	3	3		2
1l.	Guide usage	3	3		2
1m.	PIP (Picture In Picture) data of any type	1	1		2
	<i>If yes, what about these specific activities?</i>				2
1n.	Channel for PIP	1	1		2
1o.	Channel for Main Picture	1	1		2
1p.	Channel for Audio in use	1	1		2

## Data Aggregators Data Table

Num.	Please answer the following questions as completely as you can, without giving proprietary or confidential information. For questions you prefer not to answer, please check one of the two columns to the right to tell us whether we should follow-up for a response at a later time.	Yes	No	Comments	Declined to Answer, Deferred Answering,
1q.	If there are other PIP events not listed above, please describe them:	>>			2
1r.	If the STB is tuned to a mosaic channel where three or more channels are viewable on one screen, are you able to determine which channel has active audio?	2	1		4
1s.	Please list any STB activities you receive and process that are not covered in items 1a - 1r, above:  >>	<ul style="list-style-type: none"> <li>▪ Ads within guide;</li> <li>▪ interactivity, [STB] on/off, HH TV Set to STB ratio (Ed. note: not an activity, but information about the household)</li> <li>▪ interactive advertising, streaming and download services from 3<sup>rd</sup> parties such as Amazon, Netflix and YouTube</li> </ul>			2

- [Notes about Comments](#)
- Each bulleted item is the verbatim comment for a single question from a completed questionnaire.
- If a single respondent entered multiple comments, they are listed under a single bullet, separated by a comma.
- If multiple respondents made the same comment, a number in parentheses appears after the comment giving the number of respondents who made the same comment.
- If a comment was given that could identify the commenter, then the comment was either edited, if possible, or not included.

## Data Aggregators Data Table

Num.	Please answer the following questions as completely as you can, without giving proprietary or confidential information. For questions you prefer not to answer, please check one of the two columns to the right to tell us whether we should follow-up for a response at a later time.	Yes	No	Comments	Declined to Answer, Deferred Answering,
2.	<i>Please indicate whether you receive and process data for the following external devices, if connected to the STB:</i>				
2a.	DVR	0	0		1
2b.	DVD player	0	0		1
2c.	VCR	0	0		1
2d.	Game console	0	0		1
2e.	Internet Video STB (Netflix, Roku, SlingCatcher, Apple TV, Vudu, others)	0	0		1
2f.	Any other devices not listed?  >>			<ul style="list-style-type: none"> <li>▪ All STB data is from the box itself. Have relationships with Netflix, Amazon, YouTube and other 3<sup>rd</sup> party data providers with some ability to process and report on aggregate usage.</li> </ul>	1

## Data Aggregators

	Please answer the following questions as completely as you can, without giving proprietary or confidential information. For questions you prefer not to answer, please check one of the two columns to the right to tell us whether we should follow-up for a response at a later time.				Declined to Answer, Deferred Answering, or Left Answer Blank
Num.		Yes	No	Comments	

	<i>For the following questions, check the box for either <u>Yes</u> or <u>No</u>, or enter the answer in the space provided</i>				
3.	How often is raw data updated and uploaded?  <div style="border: 1px solid black; padding: 5px; width: fit-content; margin: 10px auto;">                 A number in parentheses following a comment denotes the number of respondents who entered the same comment.             </div>			<ul style="list-style-type: none"> <li>▪ Several times a day,</li> <li>▪ daily (2);</li> <li>▪ in real time;</li> <li>▪ continuously;</li> <li>▪ once per wk per STB</li> </ul>	1
3a.	Can frequency of update/upload be changed?	6	1	<ul style="list-style-type: none"> <li>▪ It's possible, but difficult</li> </ul>	0
3b.	If not, why not?  >>			<ul style="list-style-type: none"> <li>▪ Costs</li> </ul>	0
4.	What is the shortest duration tuning entry that you report (e.g., 1 sec, 5 sec, 30 sec)? _____			<ul style="list-style-type: none"> <li>▪ 1 sec (3);</li> <li>▪ 2 sec (1);</li> <li>▪ Event based, no limit generally</li> </ul>	2
4a..	Can this threshold be changed in the reporting system?	3	3		1
4b..	If not, can the threshold be changed by reprocessing of the data?	3	0		0
5.	Do you maintain your own database of program line-ups?	5	1		1
5a.	<u>If Yes</u> , from what source:  >>			<ul style="list-style-type: none"> <li>▪ Our own;</li> <li>▪ internal records, syndicated sources;</li> <li>▪ TNS;</li> <li>▪ Tribune Media Services</li> </ul>	1

## Data Aggregators Data Table

Num.	Please answer the following questions as completely as you can, without giving proprietary or confidential information. For questions you prefer not to answer, please check one of the two columns to the right to tell us whether we should follow-up for a response at a later time.			Comments	Declined to Answer, Deferred Answering, or Left Answer Blank
		Yes	No		
5b.	<u>If No</u> , do you rely on the program codes supplied in the raw data records?	1	0		0
5c.	<u>In either case</u> , is the program line-up information updated for pre-emptions and overruns?	3	1		3
5d.	Is the program line-up information adjusted for different time zones, where necessary?	4	0		3
6.	Can you identify tuning to commercial content?	5	2		0
6a.	Can you determine commercial content to the second?	5	0		0
7.	Can you differentiate national commercial content from local?	4	0		3
8.	Do you maintain your own database of channel line-ups?	4	1		3
8a.	<u>If Yes</u> , from what source: >>	<ul style="list-style-type: none"> <li>▪ Our own;</li> <li>▪ internal records, syndicated sources;</li> <li>▪ TNS</li> </ul>			1
8b.	<u>If No</u> , do you rely on the channel line-up codes supplied in the raw data records?	0	0		1
8c.	<u>Whatever your source</u> , how long does it typically take for channel line-up changes to be reflected in the tuning data? >>	<ul style="list-style-type: none"> <li>▪ They are captured the same day as users update their lineups</li> </ul>			6

## Data Aggregators Data Table

Num.	Please answer the following questions as completely as you can, without giving proprietary or confidential information. For questions you prefer not to answer, please check one of the two columns to the right to tell us whether we should follow-up for a response at a later time.			Comments	Declined to Answer, Deferred Answering, or Left Answer Blank
		Yes	No		
8d.	<u>Whatever your source</u> , can you determine the STBs able to receive a channel, so that tuning can be reported on a coverage area basis? These are often referred to as coverage area ratings.	1	3		3
9.	<p>What is the smallest geographic level for which you are able to aggregate and report tuning data? (e.g. system, county, ZIP code, block group, ad zone, etc.)</p> <p>&gt;&gt; <span style="border: 1px solid black; padding: 2px;">A number in parentheses following a comment denotes the number of respondents who entered the same comment</span></p>			<ul style="list-style-type: none"> <li>▪ We receive individual STB data, but then aggregate before reporting;</li> <li>▪ cable node;</li> <li>▪ ad zone;</li> <li>▪ HH, ZIP code (2), ZIP+4;</li> <li>▪ DMA (2)</li> </ul>	0
10.	Are non-residential subscribers' STB data combined with STB data for residential subscribers'?	4	2		1
10a.	If yes, are data for non-residential subscribers' STBs identified in such a way that users of your data can differentiate it from residential subscribers' data?	1	3		0
11.	Does seasonal occupation of households impact the percentage of STBs and Households providing usable data in certain markets?	2	1	<ul style="list-style-type: none"> <li>▪ Unknown (one answer, counted as Declined)</li> </ul>	4
11a.	If <u>Yes</u> , are seasonal households identified in the STB data?	0	2		0
11b.	If <u>Yes</u> , is the number of seasonal households reported?	0	2	<ul style="list-style-type: none"> <li>▪ Could derive based on patterns</li> </ul>	0

## Data Aggregators Data Table

Num.	Please answer the following questions as completely as you can, without giving proprietary or confidential information. For questions you prefer not to answer, please check one of the two columns to the right to tell us whether we should follow-up for a response at a later time.			Comments	Declined to Answer, Deferred Answering, or Left Answer Blank
		Yes	No		
12.	Are you able to identify tuning data by location of device within the Household? (e.g. kitchen, family room, etc.)	0	5		2
13.	Do you report characteristics of STB <i>households</i> (e.g., size of household, age of householder, presence of children, etc.)?	5	1		1
13a.	If so, how do you obtain these characteristics: >>	<ul style="list-style-type: none"> <li>▪ Modeled;</li> <li>▪ Equifax, Experian;</li> <li>▪ licensed from the sources of actual HH level data- not ascribed from geo level</li> <li>▪ a panel is surveyed quarterly for HH characteristics</li> </ul>			1
13b.	If not, do you plan in the future to report it?	0	0	▪ Possible, but not soon	1
14.	Do you report demographics of <i>each resident</i> of STB households (e.g., age, sex, education, etc.)?	3	3		1
14a.	If so, how do you obtain these characteristics: >>	▪ we currently survey age and gender of all HH residents			2
14b.	If not, do you plan in the future to report it?	0	1		2

## Data Aggregators Data Table

Num.	Please answer the following questions as completely as you can, without giving proprietary or confidential information. For questions you prefer not to answer, please check one of the two columns to the right to tell us whether we should follow-up for a response at a later time.			Comments	Declined to Answer, Deferred Answering, or Left Answer Blank
		Yes	No		
15.	Do you report demographics of <b>viewers</b> of individual programs using methods such as a survey, panel, fusion, or modeling?	1	4		2
15a.	If so, please list the method(s): >>				1
15b.	If not, do you plan in the future to use one of these methods to collect and report demographics?	0	1		3
16.	What % of STBs that <b>are enabled to</b> report data actually <b>do</b> report data each day, on average? >>	<ul style="list-style-type: none"> <li>▪ Configurable;</li> <li>▪ 80-85%;</li> <li>▪ 1/7 [one-seventh of households] are collected each day</li> </ul>			4
17.	What criteria are used to determine data usability from individual STBs? >>	<ul style="list-style-type: none"> <li>▪ Errors are used to exclude;</li> <li>▪ checks for tuning completeness;</li> <li>▪ We ensure that data is complete</li> <li>▪ we have several proprietary error checking procedures to ensure our data logs are usable</li> </ul>			3
17a.	Who determines that data is usable? We do _____ Other: _____ (check) (enter entity name)	<ul style="list-style-type: none"> <li>▪ We do (6)</li> </ul> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;">                     A number in parentheses following a comment denotes the number of respondents who entered the same comment                 </div>			1

## Data Aggregators Data Table

Num.	Please answer the following questions as completely as you can, without giving proprietary or confidential information. For questions you prefer not to answer, please check one of the two columns to the right to tell us whether we should follow-up for a response at a later time.			Comments	Declined to Answer, Deferred Answering, or Left Answer Blank
		Yes	No		
18.	How frequently are these criteria applied (e.g., daily, weekly, monthly)?  >>		<ul style="list-style-type: none"> <li>▪ Daily (5);</li> <li>▪ constantly</li> </ul>		1
19.	<p><i>Please attach a description of your data edit rules on separate sheet(s), or provide a link to a web page, indicating whether each edit is applied to the data before or after processing. Please include the following specific issues, in addition to others that are relevant to your system:</i></p> <p>Missing data</p> <p>STB on/set off</p> <p>Dwell time (the consecutive seconds tuned to a channels or event)</p> <p>Logic checks (consistency of data for each STB)</p> <p>Time Zone Adjustment (when combining data from subscribers in multiple time zones)</p>	1			6
19a.	Is there a website? (Please provide URL below)  >>	1	2		4
19b.	Is a description attached?	1	2		4

## Data Aggregators Data Table

Num.	Please answer the following questions as completely as you can, without giving proprietary or confidential information. For questions you prefer not to answer, please check one of the two columns to the right to tell us whether we should follow-up for a response at a later time.			Comments	Declined to Answer, Deferred Answering, or Left Answer Blank
		Yes	No		
20.	How do you know when an STB has malfunctioned, or stopped sending data, resulting in loss of data?  >>			<ul style="list-style-type: none"> <li>▪ Gaps in sequential event serial #s, spurious/high event counts;</li> <li>▪ disciplined series of technical and statistical tests;</li> <li>▪ we don't know</li> <li>▪ [STB] malfunctions are rare and generally limited to reboots. When a box reboots, its log is retained and continues to write once restarted. [STBs] must connect to [server] to ... be usable [viewable].</li> </ul>	3
21.	How do you determine whether lack of tuning data reflects no tuning or is due to an STB malfunction?  >>			<ul style="list-style-type: none"> <li>▪ Still refining;</li> <li>▪ depends on source and various patterns;</li> <li>▪ we don't know</li> </ul>	4
21a.	What are the types of malfunctions that can cause an STB to stop sending data, even though tuning is taking place?  >>			<ul style="list-style-type: none"> <li>▪ Retrieval errors;</li> <li>▪ platform provides technical issue logs</li> <li>▪ we do not believe this happens often enough to be material</li> </ul>	4
21b.	Are records kept of STB malfunctions?	2	1		4
21c.	How often are they updated?  >>			<ul style="list-style-type: none"> <li>▪ They are logged but can be lost if not uploaded promptly</li> <li>▪ they are part of the typical logging that is downloaded daily</li> </ul>	0
22.	Do you report statistics on STB malfunctions in your reporting and analytical system?	1	2		4

## Data Aggregators Data Table

Num.	Please answer the following questions as completely as you can, without giving proprietary or confidential information. For questions you prefer not to answer, please check one of the two columns to the right to tell us whether we should follow-up for a response at a later time.			Comments	Declined to Answer, Deferred Answering, or Left Answer Blank
		Yes	No		
23.	In your reporting and analytical systems, do you weight the raw data?	1	4	▪ Potentially	2
23a.	If so, by what variables?  >>	▪ Box type, system, MKT, age of HOH, family size, HH inc., race, Hispanic/non-Hispanic			0
24.	Do you project the data to a larger population or universe?	2	5	▪ Potentially	
24a.	If yes, how frequently are Universe Estimates updated?  >>	<div style="border: 1px solid black; padding: 5px; display: inline-block; margin-bottom: 5px;">                     One respondent gave an answer to 24a even though 24 was answered "no"                 </div> ▪ Monthly; ▪ quarterly ▪ monthly, quarterly and annually;			0
25.	If you report ratings, what Universe Estimates do you use?  >>	▪ Varies by source and intention			6
26.	Use this space to tell us any additional information about your STB measurement application or data aggregation process that you think we should know:	▪			7

**The State of Set-Top Box Viewing Data  
as of December 2009**

STB Committee of the Council for Research Excellence

**Appendices Part 2 (continued in separate documents)**

Initial Letter to Companies Invited to Participate

Data Owners Questionnaire

Data Creators Questionnaire

Data Aggregators Questionnaire

Third-party Processor Questionnaire