NIST GCR 01-818

Report on Television Data Applications

Michael A. Dolan
Technical Consultant
Television Data Broadcasting

NIST
National Institute of Standards and Technology
Technology Administration, U.S. Department of Commerce

NIST GCR 01-818

Report on Television Data Applications

Prepared for
U.S. Department of Commerce
Information Technology Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899-8900

By
Michael A. Dolan
Technical Consultant
Television Data Broadcasting

July 2001

U.S. Department of Commerce
Donald L. Evans, Secretary

National Institute of Standards and Technology
Karen H. Brown, Acting Director

Report on Television Data Applications

Prepared for the National Institute of Standards (NIST)

By Michael A Dolan

1-July-2001

Acknowledgements

The author would like to thank the studios that participated in the interviews of requirements, and to Dr. Glenn Adams for his comments on earlier drafts of this document.

1 Table of Contents

2 Terms and Acronyms

ATSC Advanced Television Systems Committee (US-centric digital television standards body)

ATVCC Analog TeleVision Closed Captioning (EIA 608-B)

ATVEF Advanced Television Enhancement Forum

Author – the generator of content, which may be from a studio, service aggregator or even the viewer.

Broadcaster – the programmer that generates a service in a local market.

Content – the aggregate video, audio and data essence and metadata that makes up the experience to the viewer, including, but not limited to, programs.

DASE DTV Application Software Environment

DBS Direct Broadcast Satellite

DDE Declarative Data Essence

DTV Digital TeleVision

DTVCC Digital TeleVision Closed Captioning (EIA 708-B)

DVB Digital Video Broadcast (European-centric digital television standards body)

EPG Electronic Program Guide

Essence – the video, audio or data elements of a program that result in some experience for the viewer (in contrast to metadata).

Metadata – the information about essence that the television distribution system generates or consumes, but does not result in a direct experience to the viewer.

MHP Multimedia Home Platform (DVB equivalent to DASE)

NABTS North American Basic Teletext Specification (EIA 516)

Program – the video, audio, data and metadata generated by a studio.

Programmer – the entity that generates a service.

Service – the serial collection of programs (also known as a channel in the US).

Service Aggregator – a cable or DBS system that aggregates services and adds value to the viewer.

SMPTE Society of Motion Picture and Television Engineers

Studio – the entity that creates content.

Viewer – the target person for the consumption of content.

3 References

[SMPTE-EBU] SMPTE/EBU, "Task Force for Harmonized Standards for the Exchange of Program Material as Bitstreams, Final Report:Analyses and Results, July 1998

[DDE] Proposed SMPTE Standard 363M, "Declarative Data Essence, Content Level 1" [1]

[DASE] ATSC Draft Standards, "DASE Introduction & Architecture", "DASE Declarative Applications", "DASE Procedural Applications", DASE Application Reference Model".

[ATVCC] EIA-608-B, "Line 21 Data Service".

[DTVCC] EIA-708-B, "Digital Television (DTV) Closed Captioning".

[NABTS] EIA-516, "North American Basic Teletext Specification"

[MPEG-VIDEO] ISO/IEC 13818-2, "INFORMATION TECHNOLOGY – GENERIC CODING OF MOVING PICTURES AND ASSOCIATED AUDIO INFORMATION: SYSTEMS

[ATSC] ATSC A/53, "ATSC Digital Television Standard"

[LATIN-1] ISO 8859, "ISO/IEC 8859: "Information technology - 8-bit single-byte coded graphic character sets".

[MRG] MRG Systems, "Teletext Tutorial", http://www.mrgsystems.co.uk/

[NIELSEN] Nielsen Media Research, http://www.nielsenmedia.com/

[1] Formerly ATVEF 1.1r26

[DATA-BOOK] McGraw-Hill, “Data Broadcasting: Understanding the ATSC Data Broadcast Standard”, Chernock, et al.

[MHP] ETSI TS 101 812, “Digital Video Broadcasting (DVB); Multimedia Home Platform (MHP) Specification 1.0”

[TVA] TV Anytime, http://www.tv-anytime.org

4 Introduction

This report characterizes various Interactive TV applications identifying those types of applications that can be implemented with:

The Internet (PC, but no Special TV Receiver)
SMPTE DDE-1 (Transport A)
SMPTE DDE-1 (Transport B) and ATSC DASE (Declarative only[2])
DVB MHP 1.0 and ATSC DASE (Procedural only)
ATSC DASE-1 and DVB MHP 1.1 (Procedural and Declarative)
OCAP

A plain TV receiver can be used by synchronizing TV and Internet content exclusively at the production end. This requires the viewer to have both a TV and a separate Internet link and browser often in separate physical devices. This is done today with Monday Night Football, and Who Wants to be a Millionaire. Related “enhancements” are provided on the Internet screen scripted, and often synchronized, to the video/audio production.

The next step up is synchronizing TV and Internet content by means of broadcasting SMPTE DDE-1 triggers using the "Transport A" (more on this later) network model.

Synchronizing TV and Internet content by means of broadcasting SMPTE DDE-1 triggers and HTML content using the "Transport B" (more on this later) network model, or with ATSC DASE Declarative content only.

Even more sophisticated applications can be done using procedural applications using Java as defined in DVB MHP 1.0 or with ATSC DASE Procedural content only.

Finally, there are systems based on both declarative and procedural environments, such as ATSC DASE-1, DVB MHP 1.1 and OCAP.

For completeness, it is worth noting that there are a few proprietary systems that have been deployed around the world with varying market penetration. These systems are not discussed further here, but for reference and future research, ones known to the author are:

AOL-TV[3]
Canal+
DTG
MHEG
OpenTV
WebTV/UltimateTV[4]

[2] Since we are discussing content here and not receiver s, this is a valid configuration of content.
[3] This is very close to SMPTE DDE-1.
[4] This is very close to SMPTE DDE-1.

Wink

This report attempts to answer the following question for each content configuration: What Interactive TV applications can the approach implement cost effectively with acceptable response times for a viewer? The characterizations include:

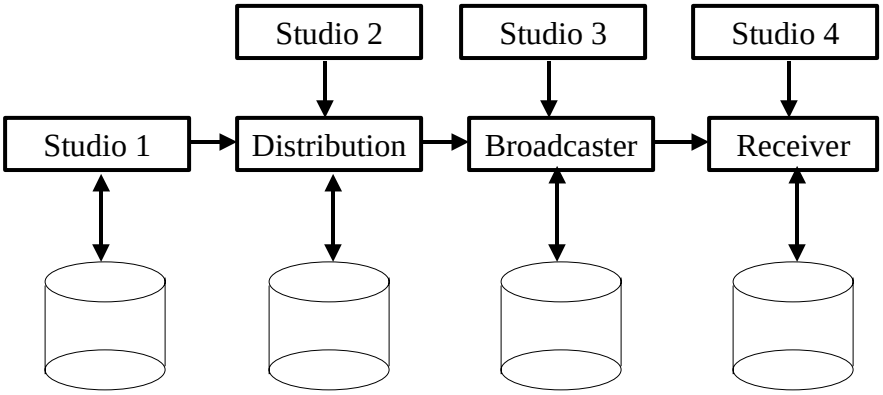
a description of the application
who would be interested in producing such an application
who is producing or planning to produce such an application
who would be the audience
which interactive TV approach can realistically implement the application
why the application requires or is best suited for a particular interactive TV approach

It is worth noting that at the time of this writing, two standards organizations are undertaking a new review of the interactive television application requirements. ATSC recently formed a new top level subcommittee, "Applications" reporting to its executive committee. And ITU, which has formed a joint rapporteurs group between the related ITU-R and ITU-T groups, called JRG-1. The reader is encouraged to keep in touch with the progress and output of these groups.

5 Production and Distribution Overview

The television production and distribution system is quite complex. A high level overview is found in Figure 5.1 below.

Figure 5.1 Television Production and Distribution



The interesting thing to note is that the production process in the traditional production studio is actually at every major step in the distribution system. That is, "content" can be created and inserted at multiple points along the way from the original creation in *Studio 1* to the *Receiver*. In fact, with the new PVR technology, the viewer has the ability to store and (soon) a certain amount of studio editing capability, adding to the complexity of the system as shown in *Studio 4*.

The requirements at each point are different. What makes sense for *Studio 1* to create may be entirely different from what *Studio 2* creates, and so on.

Further complicating things is that content can be either "live" or played from recorded media, and this status can change during its distribution. For example, a live feed from a *Studio 1* may first go to storage in the *Distribution, Broadcaster, or Receiver* to be finally played back to the viewer at a later time.

Note that the *Distribution* and *Broadcaster* are the same entity in the case of DBS and cable systems.

6 Review of Existing Data Applications

One of the major categories to consider in advancing any new technology is to ensure that at least previous applications of the old technology are accommodated. Specifically, the primary ones to consider are:

Closed Captioning
Nielsen Rating System
Teletext

These are discussed in more detail below. Note that they all are carried in the existing analog transport.

6.1 Closed Captioning

Closed captioning is traditionally thought of as an orthogonal "special" application, and it is often implemented with the minimum requirements mandated by the FCC. The original intent was to address the hearing-impaired. However, it has emerged as a useful feature for many people, often enabled by simply muting the audio on the television receiver.

Closed captioning is fundamentally a data service as everyone is thinking about it. The fact that it is treated specially in the transport is a historical artifact rather than a correct current categorization or architecture.

Closed captioning for the analog (NTSC) market is defined in [ATVCC]. However, there is already a DTV market standard defined in [DTVCC] and being implemented, so that is used for this analysis. A summary of this in the light of a data service is provided here as a requirement for future data enabled receivers. In other words, even though closed captioning is treated specially, it is easy to argue that any data environment should be able to address its requirements. Further, advanced closed captioning is likely to be implemented in the interactive environment, so it should be able to build on the basic functionality of [DTVCC].

Closed captioning is displayed in multiple rectangular windows positioned on the screen by the author. These are contained in the safe image and safe title sub-regions of the physical display.

The data service is basically a stream of 8-bit ASCII character codes from the [LATIN-1] character set. Commands and attributes are encoded and embedded in this stream along with the text itself.

The embedded commands can both define the organization and attributes of the text, as well as perform operations on previously defined windows.

The window control commands allow them to be created, deleted, and turned on and off. In addition, the manner in which the "on and off" occurs can be specified from one of "snap" (instant on/off), "fade" and "wipe" (traditional production meanings). Finally, the current window can be selected.

When a window is created, the following properties can be defined:

Color
Border
Opacity (alpha blend).

Text is added to a window according to the currently defined "pen". A pen can have the following attributes:

Size
Spacing (monospace or proportional)
"Offsetting" (direction of text)
Italic
Underline
Foreground color and opacity
Background color and opacity
Outlining
Style

0 - Default (undefined)
1 - Monospaced with serifs (similar to Courier)
2 - Proportionally spaced with serifs (similar to Times New Roman)
3 - Monospaced without serifs (similar to Helvetica Monospaced)
4 - Proportionally spaced without serifs (similar to Arial and Swiss)
5 - Casual font type (similar to Dom and Impress)
6 - Cursive font type (similar to Coronet and Marigold)
7 - Small capitals (similar to Engravers Gothic)

Metadata:

1. Dialog (normal words being spoken by characters in the programming)
2. Source or speaker ID (name of the speaker, or a description of the source of a sound)
3. Electronically reproduced voice (spoken audio heard by the characters in the drama coming from a phone, radio, PA, etc.)
4. Dialog in a language other than the drama's primary language
5. Voiceover (narration or other disembodied voice NOT heard by the characters in the drama)
6. Audible Translation (voice of a disembodied translator NOT heard by the characters in the drama)
7. Subtitle Translation (text showing a translation into the primary language of the drama)

8. Voice quality description (description of a voice quality)
9. Song Lyrics (words being sung)
10. Sound effect description (a description of a nonverbal sound or music heard by the characters in the drama)
11. Musical score description (a description of background music NOT heard by the characters in the drama)
12. Expletive (an interjectory word or expression, possible profane or harsh)
13. Text not to be displayed (reserved for future use by a text-based control and information channel within the caption text stream; e.g., hypertext, related non-caption program information)

[DTVCC] and [ATSC] define descriptors to provide specific announcement and signaling of the closed captioning (data) service. These specifications also define an encapsulation for the character stream. For example, in ATSC, it is put in a table in MPEG Picture UserData within the video encoding (see [MPEG-VIDEO]).

This is fundamentally a synchronized service as the bytes are bound to a specific frame of video and it is recommended that they be visible or acted on by the presentation of that frame. According to the specification, text and commands are transmitted ”shortly before they are to be displayed”. However, there is also a delay command which effectively permits the deferred display of a window by x seconds (1/10 second resolution) into the future.

No return channel is required for this service, as it is broadcast only with no defined viewer interaction other than on/off.

In summary, this is a synchronized, private data stream service that requires a receiver application that can decode the byte stream and display certain overlapping windows with various fonts and colors. It is closely associated with the video/audio service. No return channel is needed for this service to work. No viewer input is required, although some control over enabling and disabling the entire service is common.

6.2 Nielsen Rating System

The Nielson Rating System provides summarized viewing habits and demographics to its subscribers (typically TV Stations and Networks). It does this by collecting viewing information on selected households. Originally, this was done with written diaries, but today it is done with semi-automated boxes connected both to the television and the subscriber's phone line (the “return channel”).

The box allows individual viewers to log in and out. It also listens to the analog video signal for the rating service data that is inserted by the broadcasters. The data that is inserted (and read by the box) is:

Source Identification
Date
Time

These are encoded in 48 bits and sent in every field, replicated in both lines 20 and 22.

The source identification is a unique number that identifies the video/audio stream. It is a proprietary number assigned by Nielsen to each stream of each network/broadcaster customer. M For example, a network might have 15 source ID's, one for each feed they send out across the country. It is analogous to the MPEG source_id field used in digital transports, including both ATSC and DVB.

The date and time values are absolute date and time of day. The existing [NABTS] time service was not used, although the reason for this is not clear. It may be due to the non-absolute time reference and/or differing output from the various broadcasters. Absolute date and time of day information is supported by both ATSC and DVB digital MPEG transport tables.

The box (named the "People Meter") maintains an event log of changes to any of the tracked items - source ID, as well as individual login and logout events. It logs these along with the date and time of day at each event. Periodically (typically once per day in the middle of the night) the box sends the event log back to the Nielsen server. The Nielsen server then has to correlate this source ID and date./time information with network and broadcaster station logs in order to determine what shows were actually being watched. It also correlates the login/logout data with demographic information about each person that was collected in advance.

In summary, this is an interactive data service requiring only existing MPEG transport metadata information plus a return channel. The viewer input required is only to log in and out. It is not bound to any service, and runs at all times.

Upcoming advancements in this service include encryption of the forward source identification, and nested source identifications. The latter allows for more robust tracking of local broadcaster ad and other program insertions. Currently these have to be correlated by Nielsen from broadcaster station logs.

Clearly, the service would benefit from ISAN program identifier values being broadcast. These are unique identifiers for programs that are being standardized by SMPTE, ISO and ATSC. Along with the source ID, this would allow direct correlation with the programs being broadcast without having to correlate them with the network/broadcaster station logs.

6.3 Teletext

Teletext is an interactive broadcast data service. Like DTVCC, its broadcast encoding supports text and limited character-based graphics. Unlike DTVCC, it has the notion of multiple "pages", and usually in practice supports a return channel.

Teletext was first deployed in the UK in 1976! It has become quite popular in Europe, but never really took off at all in the US. It nevertheless is the foundation of basic interactive television data services. Teletext is currently in use in over 30 countries, with over 40 million teletext decoder chip sets shipped.

In addition to the public teletext systems, private information systems have been designed for businesses, hotels and airports. These private systems can be designed to use the full video field in which the whole TV service is used for data information, allowing a much wider range of information and/or extremely fast access times.

The primary use of teletext is for "subtitling", which includes closed captioning (just a form of subtitling in the native language). Using the pages and magazines, multi-lingual subtitling is possible and now becoming more popular. Other simple applications include program guides and VCR recording control. More complex applications in use are:

- Real Time Stock Quotes
- News Services, including pictures and graphics
- Travel Information:
  - Airline and Rail Schedules
  - Hotel Availability
  - Road Conditions
  - Weather Conditions
- Local Emergency Information

A teletext service consists of up to eight magazines. Each magazine consists of up to 100 pages. The magazine number is implicitly encoded in the page number. Each page also has an associated sub-page which can be used to extend the number of individual pages within each magazine.

A teletext service will broadcast many pages on a carousel, and the viewer can request any specific page. Typical carousel times are around 30 seconds, so even in receivers that cannot store the pages in advance, access time averages around 15 seconds for each page.

The character and graphics matrix is 40 across by 24 down. While this is a fairly crude by today's web graphics standards, it still provides a a very viewable basic graphics service.

Some metadata information that may be broadcast along with the text/graphics itself is:

- Network Identifier
- Local Channel Identifier
- Programme Identifier
- Country
- Date & Time

The encoding in use permits 45 bytes in one line. Many broadcasters allocate 8 lines per field for teletext, for a total of 360 bytes per field. One page of data requires 24 packets (one packet per character line), or about 3 total fields. This results in about 17 pages per second on most systems concurrently carrying video and audio as well.

7 Authoring Requirements & Considerations

7.1 Introduction

In order to better understand the requirements of "interactive television", the author interviewed a variety of people in the industry. Several people in each category of the distribution provided their thoughts about what should be done or what they wanted to do with the upcoming technology. Those interviewed came from traditional programming studios, broadcasters, and cable and DBS head-ends. All of which are "authors" of interactive television content.

The discussion and requirements in this section are a composite of information obtained through all the interviews most of whom wished their contributions to remain anonymous. The reader is cautioned about drawing any conclusions about the source of any specific statements made here.

The first thing that needs to be pointed out is that, without exception, the input about applications was always clouded by the knowledge of what basic system level technology was available to the author. In other words, it was not possible to separate the pure authoring requirements and desires of the creative talent from the context of what was reasonable given the direction of the industry in receiver architecture and technology. In discussions with the more creative talent, there was usually a gap in what they could even imagine doing. And, as one talks to more technically-inclined authors, their thinking was targeted towards what they knew to be currently (or available shortly) possible. This results in a circular dependency between the technology that is available and the possible applications that may be inspired.

It was also very clear that there is a lack of clear business models and that the industry is in a highly experimental mode of trying things to see what works. This includes employing focus groups, and other testing, marketing and promotion. These activities are not only determined to find the business models that work, but also to see what technology that is being proposed left and right by suppliers will actually work for them.

So, the end result is that the applications requirements review is a complex and iterative circular process of thinking about what is desirable in the context of what is technically possible. This circular process will continue for some time to come before "interactive television" is fully defined and realized for the average viewer.

7.2 Summary of the Application Scenarios

During the course of all the interviews and discussions, there were a set of application scenarios that came forth. These are summarized below and described briefly in this section. There is no attempt here to try and prioritize these or assess their viability, either long term or short term. They should be considered a reasonably complete list of categories of applications that are likely to be experimented with.

Coupons
Electronic Program Guide (EPG)
EPG Interaction
Gambling
Gaming
Home Shopping
Inquiry
Look and Feel
Mark for Recording
Polling
Sports Statistics
Train/Plane Schedules
Targeted Advertising
Viewer Demographics
Weather

7.2.1 Coupons

This is the offer for a discount or other incentive to the viewer for a product or service. This could range from the discounted pay per view movies from the broadcaster or aggregator itself, or for items currently being advertised. This is different from "inquiry" below in that it could be done without any interaction by the viewer. This has the interesting receiver requirement that there be some sort of output device to either print the coupon or otherwise transfer its value to the viewer.

7.2.2 Electronic Program Guide (EPG)

This is the display on the receiver of the services and programming that is coming up. This data is carried either in the television signal or out of band, and does not require any sort of return channel (see more below on enhanced variations on this). Also, the interaction is local to the receiver – that is, it is used primarily for tuning or recording or other things the viewer does locally to the receiver.

7.2.3 EPG Interaction

This is a richer EPG experience that involves retrieving program reviews and other activities that require a return channel. Other examples include: purchasing pay per view movies, purchasing DVD's, purchasing audio CD's, obtaining more information about the movie or audio program such as artists, directors, etc.

7.2.4 Gambling[5]

This is the ability to wager money for all the usual sorts of gambling scenarios, such as:

Racing (horse, dog, auto, etc)
Sports (football, soccer, basketball, golf, etc)
Individual Games of Chance (poker, 21, etc)

Betting on live events such as sports and races provides the opportunity to offer this data enhancement in conjunction with the video/audio of the event itself. This is a more compelling for the television broadcast than simply a game without any video/audio context.

Even games of chance, such as poker or 21 card games could be played using a full-time return channel to provide the necessary real-time interaction.

7.2.5 Gaming

This is the live interaction with a game show where the viewers play along with the game show contestants and optionally each other. Prizes may be awarded for high scores. Detailed statistics about the answers are provided, etc. This requires a robust return channel and central processing computers. It also has limitations when not done in real time or not done live worldwide.

7.2.6 Home Shopping

This is the traditional service provided by the Home Shopping Network, QVC, etc. The only difference is that instead of having to go out of band with a voice call to an 800 number or go to one's PC to visit a website, the products can be purchased with a single button action by the viewer. This requires a business model in place and some (although modest) return channel functionality.

7.2.7 Inquiry

This is a simple action by the viewer to request more information about the current programming. It would most often be used in advertising to send out product literature. The difference between this and "coupons" above is that this is initiated by the viewer and may or may not have any prompting in the program.

[5] At authoring time, the State of Nevada just approved online betting. While some are concerned about the government approval of this application for broadcast television, it is the author's view that its application, perhaps constrained to cable networks, is only a matter of time.

7.2.8 Look and Feel

This is the ability of the (usually receiver manufacturer or aggregator) to control the viewer experience while outside the programming, such as when viewing an EPG. Traditionally, this is done by the receiver manufacturer, however the aggregators are exerting more interest in branding and offering enhanced navigation features.

7.2.9 Mark for Recording

This is ability of a viewer to mark a program for recording. This is a basic feature of the new PVR receivers, but it is an important interactive application. The fact that it is a local interaction with the receiver is interesting technically, but does not make it any less important. This function usually includes a range of sophistication in managing what is being recorded, the storage systems, etc. This is just an extension of the VCR recording of the past.

7.2.10 Polling

Polling is the ability of the viewer to register an opinion, usually in the form of a yes/no vote. There is quite a bit of programming (mainly from local and network news) that does this now. The only difference is that instead of having to call a special phone number or go to the broadcaster's web site, a simple one-button-push is done by the viewer without leaving the couch.

7.2.11 Sports Statistics

This is the live action enhancement of a sports event such as football, where the viewer is presented with the option of "browsing" various statistics about the game, the individual players, past games, etc. This can be done on a single screen, but is much more robust on 2 screen (either virtual or physical). The statistics are both scrolled live, as well as carouseled during the program.

7.2.12 Train/Plane Schedules

This is a popular teletext application today (although not in the US). Pages of schedule data are carouseled during the broadcast of video and audio programming, or in some cases on a dedicated service. The viewer can navigate locally through the data pages. In more advanced versions of this, by using a return channel, tickets could be purchased.

7.2.13 Targeted Advertising

Today, advertising is pretty constant for all viewers, except for local ad insertion by broadcasters. However, when a receiver has enough demographic information stored in it, and enough forward service bandwidth is available, advertising can be more highly targeted to the viewer. For example, if the receiver knew the basic range of household income, and automobile ad would select the model of car most likely to appeal to that household (e.g. a Cadillac versus a Geo ad). Another example of this is an ad whose video and audio is the same for everyone, but has a text overlay that shows the nearest retail store to the viewer based on the viewer's zipcode.

7.2.14 Viewer Demographics

This is a more advanced version of the Nielsen Rating system. First, some amount of this is needed to implement the targeted advertising above. But more importantly, this can provide valuable (to the broadcaster and programmer) statistics about what is being viewed and when, along with basic household information. Obviously such active feedback generates privacy concerns that have to be managed.

7.2.15 Weather Information

This is an example of a generic category of programming for which the primary information being conveyed is data. Weather is just the most obvious example. This is the ability to complement the video/audio programming with summary information that can be called up by the viewer on demand. In the weather example, the viewer could bring up the local forecast without waiting for the video/audio carousel to come around again.

7.2.16 Other

In some interviews, there was a general theme of trying to integrate traditional Internet experiences with television viewing. Examples of this are email, web browsing as well as some proprietary system features. While this may or may not be a viable receiver product integration step, it was omitted from this study since it did not offer any new viewer experience, but rather an integration of existing experiences. This paper focuses on new applications that are used in conjunction with the video and audio programming.

Not interviewed, but certainly in existence today, are "data services" that are unrelated to the video and audio. These are simply using the digital television signal as a transport for the carriage of a variety of services that could be delivered by other means as well, such as Internet, satellite, etc. Again, since these are not television programming applications (i.e. not related to any video and audio), these too are omitted from the discussion.

Finally, notably absent from the interviews (in the author's opinion) was pornography. This may be due to the aversion to it from traditional family programming, ratings phobia and FCC regulations today. However, pornography has traditionally launched the initial business models for many new technologies. And, the author would expect this to play an important role in

interactive television. It may be constrained to cable environments, and almost surely requires a CA system to properly control and deploy. But an analysis of likely television data applications would not be complete without at least mentioning this.

7.3 Types of Authors and Their Primary Applications

As shown in Figure 5.1, there are different kinds of authors in the television distribution system. There are those in the studios that are working to enhance the basic programming by adding data to the video and audio they traditionally produce. The location of the studio can be anywhere in the distribution. It is characterized by the need for a data-capable edit station. And, there are aggregators such as cable systems and DBS systems. These authors are less interested in enhancing any particular program, but rather offering wider services. Finally, there is the receiver itself which offers some applications to the viewer independent of the programming being broadcast.

The applications summarized above can be categorized by type of author, or source. This is helpful to better understand what tools may be needed at what point in the distribution and what receiver capabilities may be required in different configurations (for example, terrestrial broadcast versus cable/DBS systems). Table 7.1 below provides this summary. Note that all applications are arguably in all source categories, but this comes from where the application interest lay in the interviews.

Table 7.1 Sources of Applications

Application	Programmer	Aggregator	Receiver
Coupons	X	X	
EPG		X	X
EPG Interaction	X	X	X
Gambling	X		
Gaming	X		
Home Shopping	X		
Inquiry	X	X	
Look and Feel		X	
Mark for Recording			X
Polling	X	X	
Sports Statistics	X		
Train/Plane Schedules	X		
Targeted Advertising	X	X	
Viewer Demographics	X	X	
Weather	X		

7.4 Authoring Considerations

There are several system level technology factors in considering the type of applications that are possible. While normally it would be better to discuss these as part of the solution space, they affected the authoring discussions above, so they deserve some discussion here. The main system factors are:

Number of Screens
Connectivity
Financial Model
Broadcast Timing
Cost of Production

7.4.1 Number of Screens

The ITV experience has been evolving over time to a convergence in some manner between the traditional television video/audio presentation and some form of related "data" content. This is made available through some sort of distribution (more below on connectivity). In some cases, the related material is displayed on the same screen as the video and audio. A teaser to purchase a product is an example. Alternatively, an application may be better displayed on a 2nd screen separate from the video. This latter form has been called "the 2-screen scenario". Note that this can be implemented as two physical screens or the ability to switch between 2 virtual screens on the same physical display device. And, this may be a computer screen, an eBook screen, etc.

A single screen generally provides poor navigation, and is capable only of "light" interaction. This is the result of two main factors:

Distraction from the main overall purpose of the programming
Best use of screen real estate

A summary of the applications and which need which mode is summarized in Table 7.2 below (author's opinion). In some cases, the applications could be done in more than one mode depending on how it was done, so there are a few that have more than one column marked. Also, note that there are a few applications that require no screen at all. The signal to the viewer is implied in the video/audio, or none is truly required.

Table 7.2 Screen Requirements of Applications

Application	No Screen	One Screen	Two Screens
Coupons	X		
EPG			X
EPG Interaction			X
Gambling		X	X
Gaming		X	X

Home Shopping	X	X	
Inquiry	X	X	
Look and Feel			X
Mark for Recording			X
Polling		X	
Sports Statistics		X	X
Train/Plane Schedules			X
Targeted Advertising		X	
Viewer Demographics	X		
Weather		X	X

One measure of the viewer experience complexity is how many screen are needed to accomplish the application. This will affect the production cost as well as the type of involvement of the viewer.

7.4.2 Connectivity

The connectivity of the receivers plays a large role in the applications it is capable of handling. The connectivity includes the forward channel as well as the return channel. These can take the form of a single broadcast television signal (with data), a part-time proprietary return channel, a full-time 2-way Internet connection, or some combination of all of them. The related data, in addition to being delivered in the broadcast, could also be delivered outside the broadcast stream over an Internet connection for example. This might be done with the volume of data is significant or the data is highly targeted to individual viewers, thus requiring a dedicated connection.

Adding a simple, part-time return channel, it is also possible to do all the "light" interaction applications listed above. Many of these can be done with a deferred callback, hence the "light" return channel notion.

With a "serious" forward and return channel connection (such as DSL or cable-modem and/or digital forward data channel), then all the data and interactive-intense applications are possible. Many authors are assuming this richest connectivity mode will be in place, and that it will not incur any additional cost to use it for enhancing the programming.

Table 7.3 Return Channel Requirements of Applications

Application	None	Part Time	Internet
Coupons	X		
EPG	X		
EPG Interaction			X
Gambling			X
Gaming			X
Home Shopping		X	
Inquiry		X	

Look and Feel	X		
Mark for Recording	X		
Polling			X
Sports Statistics	X		
Train/Plane Schedules	X	X	
Targeted Advertising	X		
Viewer Demographics	X		
Weather	X		

As in the number of screens, the type of return channel required by an application may affect its production cost (will require server systems in addition to more expensive receivers).

7.4.3 Financial Model

There are two fundamental financial models that affect the applications. The first and foremost is anything that increases viewership. However, there is also a new model – direct financial transactions. The former is the traditional goal of any new production technique – increase the number of people watching the programming (ratings). The latter is a new model that is being tried where viewers are given the opportunity to buy something or otherwise directly cause a financial event to occur. Finally, there is the ability to provide targeted advertising.

The increased viewership can come from providing extra information about the video/audio programming – making the experience richer. This could be in the form of sports statistics or technical details on a topic in the broadcast, etc. The author's opinion of the financial model categorization is summarized in Table 7.4 below.

Table 7.4 Financial Model Requirements of Applications

Application	Unknown	Increased Viewership	Direct Revenue
Coupons			X
EPG	X		
EPG Interaction		X	
Gambling		X	X
Gaming		X	
Home Shopping		X	X
Inquiry			X
Look and Feel	X		
Mark for Recording	X		
Polling		X	
Sports Statistics		X	
Train/Plane Schedules		X	X
Targeted Advertising			X
Viewer Demographics	X		
Weather		X	

7.4.4 Broadcast Timing

There are 3 broadcast scenarios that affect how applications will work that I've named, "timing" of the broadcast. They are:

Live
Taped
Syndicated

When doing interactive applications, these scenarios greatly affect the type and nature of the applications that are possible. For example, it may be difficult or impossible to do a game show on syndicated programming since it can be broadcast at random dates and times in every market. Even a taped show, or one that is delayed for time zones, can pose some challenges and practical limitations. Similarly a live broadcast has its own challenges in the authoring (has to be done in real time) as well as response system behavior if the broadcast reaches a large international audience, such as the Super Bowl. The possible affect of broadcast timing on applications is shown in Table 7.5 below.

Table 7.5 Broadcast Timing Requirements of Applications

Application	Live	Taped	Syndicated
Coupons	X	X	X
EPG	X	X	X
EPG Interaction	X	X	X
Gambling	X		
Gaming	X	X	
Home Shopping	X	X	
Inquiry	X	X	X
Look and Feel	X	X	X
Mark for Recording	X	X	X
Polling	X		
Sports Statistics	X		
Train/Plane Schedules	X	X	
Targeted Advertising	X	X	X
Viewer Demographics	X	X	X
Weather	X		

7.4.5 Cost of Production

The cost of production will affect the success of an application idea. This is just basic television production 101. But some applications have hidden costs such as server farms and real-time costs all of which can be quite high and beyond the authoring of the data itself. And, any real-time authoring is going to be difficult, just as it is for video and audio. Table 7.6 below shows the author's estimate of the relative costs of the applications. Note that the "cost" is per program, so in the case of receiver functions, this will bias them to be low since the costs are amortized over many services and programs over time.

Table 7.6 Estimated Cost of Production of Applications

Application	Small	Modest	High
Coupons	X		
EPG		X	
EPG Interaction		X	
Gambling			X
Gaming			X
Home Shopping		X	
Inquiry	X		
Look and Feel	X		
Mark for Recording	X		
Polling		X	
Sports Statistics		X	
Train/Plane Schedules		X	
Targeted Advertising	X		
Viewer Demographics	X		
Weather		X	

8 Overview of Content Architectures

There are a variety of architectures and technology being proposed for use in interactive television and processing of data essence. In most cases the definition of the architectures was driven by available technology. These have been standardized, or are in the process of being standardized, by various standards bodies in the world. It is interesting to note that in all cases, while content authors have usually participated at some level, the design work was primarily driven by receiver platform manufacturers. Hence it will be more convenient to describe the architectures by reference to the receiver model, rather than an authoring model.

There are several systems we will review here and analyze with respect to the requirements of the applications described in the preceding sections. These architectures are:

None (separate Internet browser)
SMPTE DDE-1 (Transport A)
SMPTE DDE-1 (Transport B) & ATSC DASE 1.0 (declarative portion only)
DVP MHP 1.0 & ATSC DASE-1 (procedural portion only)
DVB MHP 1.1 & ATSC DASE-1
OCAP

These are each discussed in more detail below.

8.1 None (separate Internet browser)

This architecture is one where the television broadcast and television receiver have no special capability with respect to data essence. The television shows are being enhanced in some manner through the exclusive use of the Internet and a web-based computer. “Normal” viewers are not provided any extra experience. The extra experience comes from a household computer (typically a Windows PC) with web browser. The experience can be a passive one from the author, requiring the viewer to navigate around on his own, presumably in parallel with the video/audio broadcast.

Alternatively, through the use of scripting, the author can push” an data experience to the viewer, and optionally provide some loose synchronization with the video/audio programming.

Finally, the single household computer could also have a video receiver built in, which could provide the video/audio experience on the same screen as the web browsing experience. The only connection between the two is a compositing by the household computer.

There are no standards for this experience other than the loosely implemented web browser standards. Content is often authored with a lot of conditional scripting to broaden the receiver platforms it supports.

A full time Internet connection (return channel) is required.

8.2 SMPTE DDE-1 (Transport A)

SMPTE is very close to completing the standardization of a declarative-based architecture, known as Declarative Data Essence, Content Level 1 (DDE-1). This includes the common web browser technology:

HTML 4.0 (All 3 Document Type Definitions)
CSS1
ECMAScript
DOM-0

Additionally, it includes decoders for the following content types, which may be embedded in the HTML:

ASCII Text (text/plain)
PNG Images (image/png)
JPEG Images (image/jpeg)
PCM Audio (audio/basic)

URI scheme support includes:

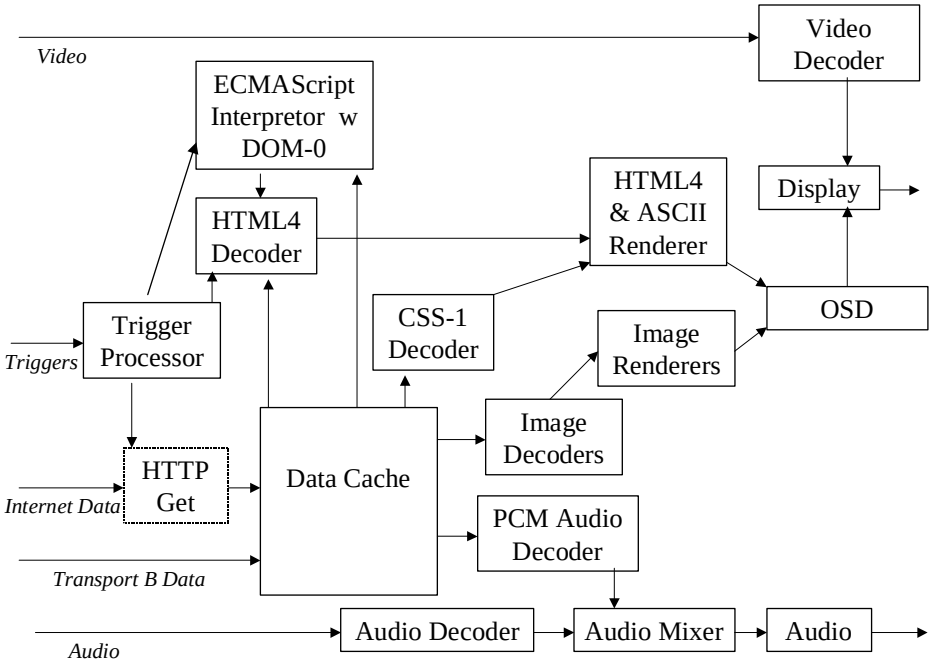
tv:
lid:
http:

And, finally, it includes support for a "trigger". A trigger is a special data type that is carried in the video/audio stream and provides synchronization between the DDE-1 content and the video/audio programming. The synchronization is a bit loose since the system assumes zero decode time for both the trigger and the target content. The trigger is typically a URI (http: for Transport A).

"Transport A" is an architecture where the triggers are sent in the video/audio broadcast, but the rest of the content is actively retrieved by the receiver from the Internet in response to the triggers. Like the "none" case above, a full time Internet connection (return channel) is presumed and required for any application to work. For more information, please see [DDE].

A typical DDE-1 receiver is shown in Figure 8.1 below.

Figure 8.1 Typical DDE-1 Receiver.



8.3 SMPTE DDE-1 (Transport B) & ATSC DASE-1 (Declarative Only[6])

The "Transport B" version of DDE-1 delivers the data over the video/audio broadcast along with the triggers. The Internet return-channel in this architecture is optional, since all the data essence could be delivered in the video/audio broadcast. The triggers are more likely to use the lid: URI scheme, but can use both lid: and http:.

ATSC is very close to completing the standardization[7] of its DTV Application Software Environment (DASE). It includes two main application environments: Declarative and Procedural. For this architecture we will constrain it to only a declarative application (DA) only. Due to the primary functional requirement of the DASE-DA, it is functionally very similar to SMPTE DDE-1, the only exception being no return-channel is defined.

ATSC-DA is comprised of the following Web technology:

XDML (XHTML1.0 Traditional and Frameset subsets)
CSS2 Subset
ECMAScript
DOM2 Subset

[6] This is not meant to characterize a possible DASE receiver as "declarative only", but rather define it through an authoring constraint.
[7] This document is based on the ATSC work in process as of this writing. The reader is cautioned that the final DASE-1 Standard may differ from this and final publication from ATSC should be consulted.

Additionally, it includes decoders for the following content types, which may be embedded in the XHTML:

PNG Images (image/png)
JPEG Images (image/jpeg)
PCM Audio (audio/basic)
MNG Animated Graphics (video/mng)
Portable Font Resource (TrueDoc®) Fonts (application/font-tdpfr)
Zip Archive (application/zip)

URI scheme support includes:

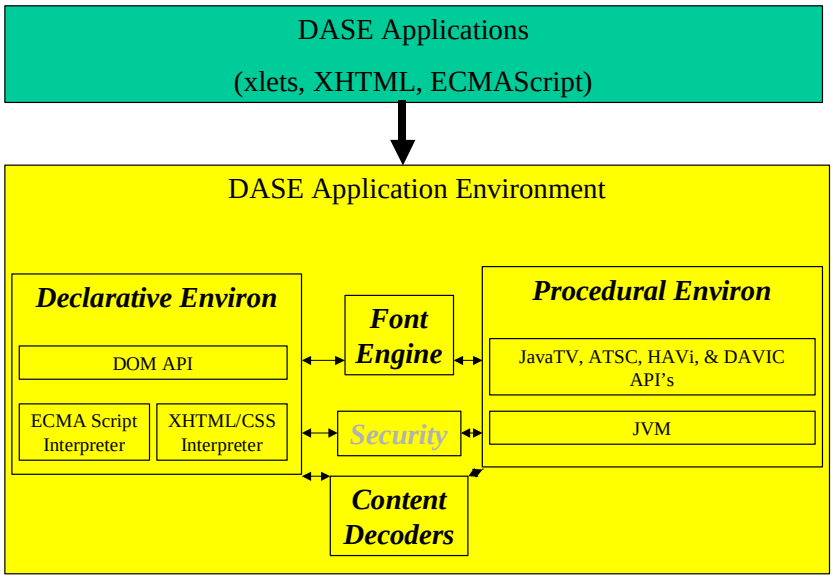
tv:
lid:

And, finally, it includes support for a "trigger". A trigger is a special data type that provides synchronization between the DASE-DA content and the video/audio programming. The synchronization is currently loose since the system assumes zero decode time for both the trigger and the target content. The trigger is an event as defined at W3C. This is a richer, more generalized version of the DDE-1 trigger.

The main functional additions in DASE-DA are the support for MNG video files and downloadable fonts.

There is no full time Internet connection or return channel. For more information, please see [DASE]. A block diagram of a typical ATSC DASE receiver is shown in Figure 8.2 below. Note that the DA portion is all we are focused on in this section (ignore the PA environment).

Figure 8.2 Typical ATSC DASE-1 Receiver.



8.4 DVB MHP 1.0 & ATSC DASE-1 (Procedural Only)

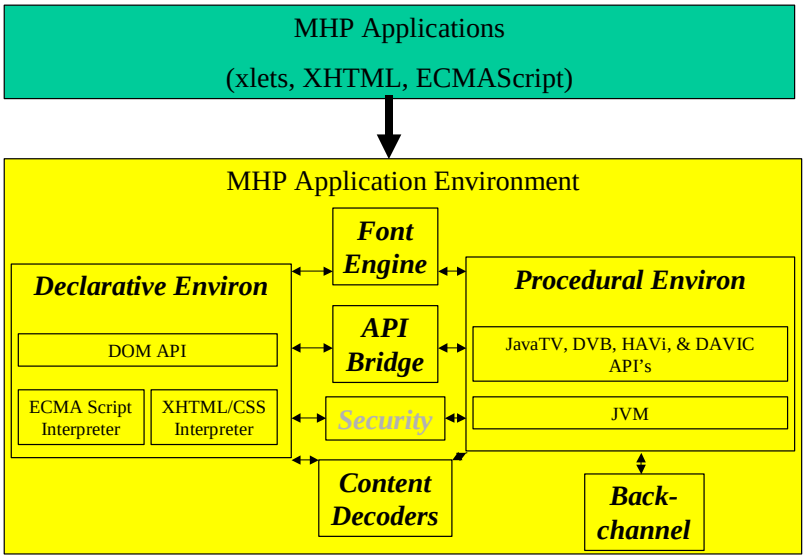
Both DVB MHP 1.0 and ATSC DASE-1 define a procedural environment based on the Java Virtual Machine and a set of special API's. The two environments are very similar in that the API support is based around:

Personal Java
Java Media Framework (JMF)
Java® TV
HAVi

In addition, they also each contain some private API's (org.atsc.* and org.dvb.*) for support of the differences between the transports as well as functions not standardized in Java TV. So, while it may be possible to author content that works on both environments, it is likely that the behavior will vary or environment-dependent classes needed/used.

See Figure 8.2 for a general picture of the architecture of the entire DASE environment. Note that we are focused only on the PA environment in this section. See Figure 8.3 below for a picture of a typical MHP 1.1 receiver. For the procedural only environment, ignore the left half of the diagram.

Figure 8.3 Typical DVP MHP 1.1 Receiver.



8.5 ATSC DASE-1 & MHP 1.1 (Procedural and Declarative)

Finally, by combining the Declarative and Procedural environments, one has the union of the architectures described above. See Figure 8.2 for the entire DASE-1 environment, and see Figure 8.3 for the entire MHP 1.1 environment, both of which have a similar architecture and design. The major function difference is that DVB MHP 1.1 contains a return-channel.

8.6 OCAP

A standard for the US cable industry is underway[8] in the OpenCable group managed by Cablelabs, called OCAP. It is an environment that is basically the union of SMPTE DDE-1 Transport B and DVB MHP 1.0.

8.7 Summary of Platform Capabilities

Table 8.1 shows a summary comparison of the platform capabilities important to authoring.

Table 8.1 Summary of Platform Capabilities.

Capability	None	DDE-1-A	DDE-1-B DASE-	DASE-PA	MHP 1.0	MHP 1.1	OCAP

[8] This information about OCAP is based on the presentation made by Cablelabs at the NIST Symposium held on 21-June-01.

			DA				
Markup Languages	X	X	X			X	X
ECMAScript	X	X	X			X	X
DOM	X	X	X	X		X	X
Image Content	X	X	X	X	X	X	X
PCM Audio	X	X	X	X	X	X	X
Java Virtual Machine	X			X	X	X	X
Java TV API's				X	X	X	X
HAVi API's				X	X	X	X
JMF API's				X	X	X	X
Return channel		X			X	X	X

9 Suitability of Each Receiver Configuration to the Application Requirements

9.1 None (separate Internet browser)

While the least defined, this is the most powerful of the environments. It relies on household computer capability which has grown to be quite rich. And, often a Windows® and Pentium® platform is assumed where the content authors actually include executables and plugins for this specific platform. This leads to a very open ended capability set, and thus, all applications **could** execute on this platform. However, it forces a physical dual-screen environment, and it is not possible to have reliably synchronized applications. The applications that seem workable in some reasonable manner are:

Teletext
Gambling
Gaming
Polling
Train/Plain Schedules
Weather

9.2 SMPTE DDE-1 (Transport A)

Transport A version of DDE-1 requires a return channel to work. Hence it has a few extra things it can do that Transport B cannot reliably do. It is declarative only, so applications that require general purpose computing cannot be done. Unlike "none", it uses a single screen so advertising and other applications requiring coupling to video can be accomplished. The applications that seem workable in some reasonable manner are:

Teletext
Coupons
Gambling
Gaming
Home Shopping
Inquiry
Polling
Plane/Train Schedules
Weather

9.3 SMPTE DDE-1 (Transport B) & ATSC DASE 1.0 (Declarative Only)

Transport B DDE-1 removes the requirement for a return channel. Thus it has a reduced application set with respect to Transport A. Since DASE DA is functionally very similar (the extra functions do not increase application support), then the list is identical. The applications that seem workable in some reasonable manner are:

Teletext
Coupons
Look and Feel
Train/Plane Schedules
Weather

9.4 DVB MHP 1.0 & ATSC DASE-1 (Procedural Only)

ATSC DASE PA provides a robust general-purpose computing environment, but it lacks a return channel. Hence many applications that it could otherwise do are not possible. But, the applications that seem workable in some reasonable manner on ATSC DASE 1.0 PA are:

Coupons
EPG
Look and Feel
Train/Plane Schedules
Weather

In contrast, the MHP 1.0 environment can do **all applications** except for "Mark for Recording". The PVR functions like this are relatively new, else it would probably have that as well. The primary reasons for this environment being able to handle the applications are:

Return channel
General Purpose Computing

9.5 ATSC DASE-1 (Procedural and Declarative)

This is just the union of the DA and PA capabilities, and thus the supported applications. The union list is:

Teletext
Coupons
EPG
Look and Feel
Train/Plane Schedules
Weather

9.6 DVB MHP 1.1 (Procedural and Declarative)

This is MHP 1.0 (including the return channel) and a DA environment similar to DASE-1, hence is supports all the application scenarios.

9.7 OCAP (Procedural and Declarative)

This is DDE-1-B, plus MHP 1.0. Since DDE-1-B is functionally almost equivalent to DASE-1 PA and MHP 1.1 DA, it is therefore functionally equivalent to MHP 1.1, and thus supports all application scenarios.

9.8 Summary of Application Analysis

Table 9.1 below shows a summary of all the applications discussed in this report and which are likely to work or are best suited to each environment. As a general rule, **anything** can be done in a procedural environment with a return-channel. So, you will find the MHP columns and the OCAP column to be complete. However, the question of authoring cost affects this utility. For example, what may be a suitable authoring cost using HTML may be excessive using Java. Thus, the fact that one **could** do anything in Java TV is only one factor in considering what is best. This weighs in favor of the DA only environment (DDE).

The second driving factor (not addressed in this "authoring" report) is receiver cost. Clearly, it is more expensive to implement both the DA and PA environments than only one environment. So, without making any claims as to which environment may be more costly to implement in a receiver, clearly authoring both is more expensive than authoring one. So, this weighs in favor of the DA or PA only environments (DDE, DASE PA, MHP 1.0).

It is the author's opinion that in the near term, authoring complexity (i.e. production cost – see Section 7.4.5), and receiver cost will both be large factors in adoption of specific ITV environments.

The other thing of note is that the "Mark for Record" applications are not directly supported in any environment (although MHP comes close, there is currently no content identification to make it actually usable). This is one of a class of applications grouped into "personal video recorders" or PVR's, and are being addressed at this writing by the TV Anytime group [TVA]. While this is primarily a receiver functional set, it is an important emerging capability that will, in the near term, overshadow the advanced environments discussed here, in the author's opinion. The reader is encouraged to follow this work closely, particularly as it works through the content rights management issues.

Table 9.1 Summary of Application Suitability to Platforms.

Application	None	DDE-1-A	DDE-1-B DASE-DA	DASE-PA	MHP 1.0	MHP 1.1	OCAP
Closed Captioning				X	X	X	X
Nielsen Ratings					X	X	X
Teletext	X	X	X	X	X	X	X
Coupons		X	X	X	X	X	X
EPG				X	X	X	X
EPG Interaction		X			X	X	X
Gambling	X	X			X	X	X
Gaming	X	X			X	X	X
Home Shopping		X			X	X	X
Inquiry		X			X	X	X
Look and Feel			X	X	X	X	X
Mark for Recording							
Polling	X	X			X	X	X
Sports Statistics	X	X	X	X	X	X	X
Train/Plane Schedules	X	X	X	X	X	X	X
Targeted Advertising	X			X	X	X	X
Viewer Demographics					X	X	X
Weather	X	X	X	X	X	X	X