On track

A primer on media asset identification

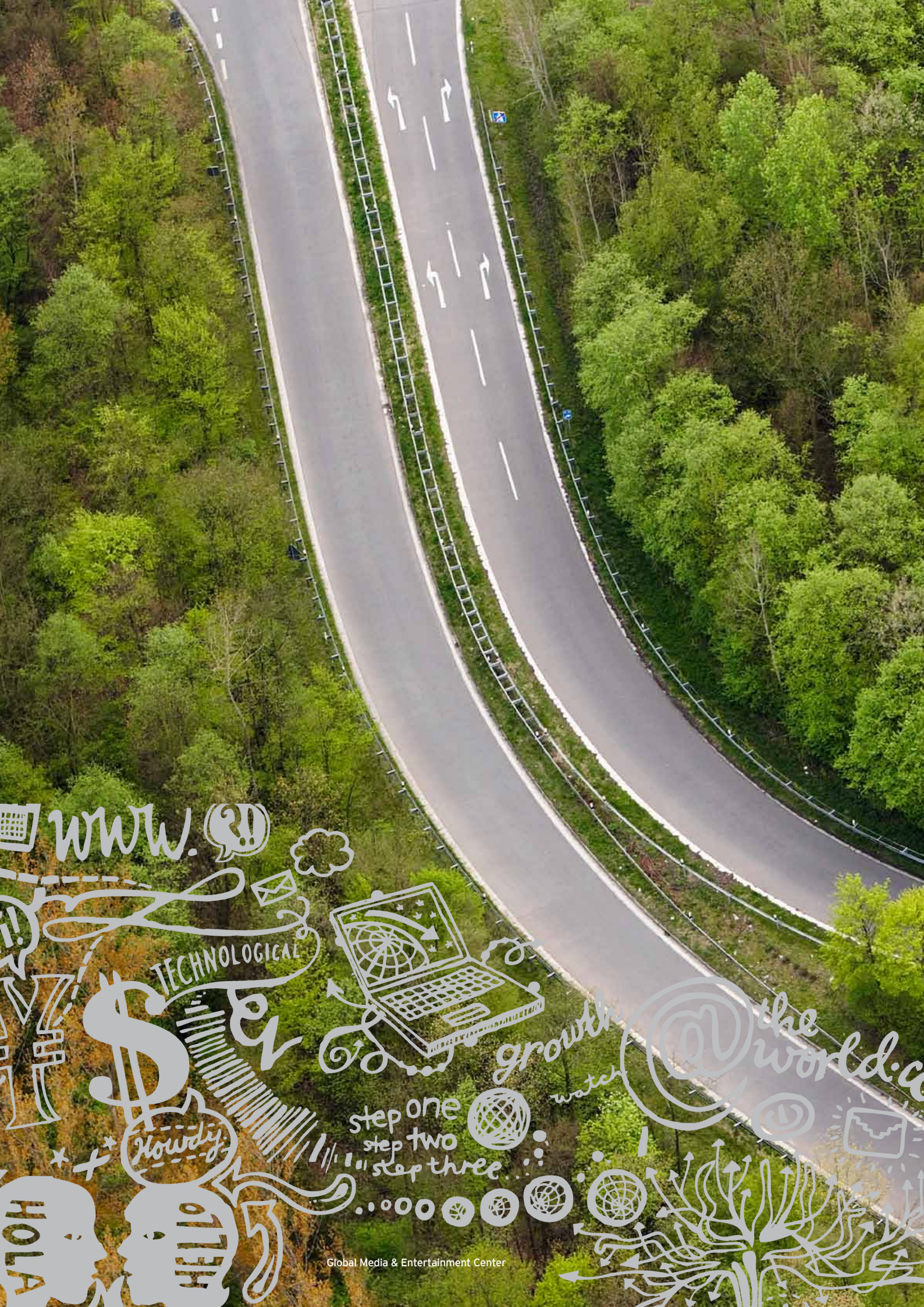
May 2011

Ernst & Young
Quality In Everything We Do

cimm
coalition for innovative media measurement

Conten

ts



Global Media & Entertainment Center

1.0 Introduction

The media and entertainment (M&E) sector is fundamentally transforming. Rapid technology development has given rise to several new platforms, and consumers are eagerly embracing them. M&E companies are responding to this change by shifting their content and distribution strategies. They are developing new products and services to meet evolving consumer demands.

With this shift in mind, content owners and advertisers are producing more content than ever, but they are finding it increasingly difficult to monetize these assets. The entire M&E supply chain is finding it harder to track where and how often their assets are consumed. Without this fundamental building block of media transparency and accountability, content owners will be challenged to fully monetize their entertainment assets and advertisers will be faced with a less and less effective means for matching advertising, content and audiences.

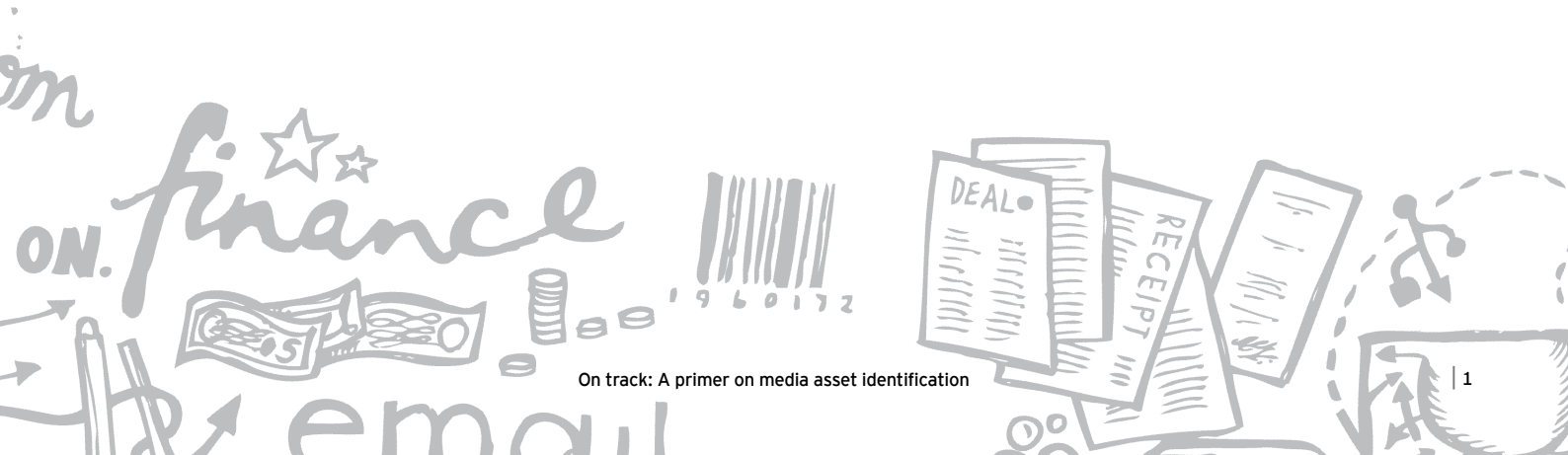
There are several M&E industry trade groups and coalitions wrestling with how to improve media transparency and accountability. One such organization is the Coalition for Innovative Media Measurement (CIMM), whose *Trackable Asset, Cross-Platform Identification initiative* (TAXI), is designed to establish open and interoperable standards upon which incumbent business applications and supporting operational processes can more effectively adapt to the requirements of asset tracking. Standardized, cross-platform asset identification can simplify a plethora of business, technical and operational challenges. It has become imperative in monetizing assets in an increasingly complicated and ever-changing media and entertainment value chain.

What to expect in the pages that follow

Ernst & Young (EY) developed this asset identification primer while working with CIMM on its TAXI Feasibility Study. This primer is designed to help industry executives better understand the language of asset identification and tracking, and help them gain a perspective on key ID methodologies available to M&E ecosystem participants.

In this primer, EY lists many significant asset identification registries and ID schemata available to content and advertising companies that participate in the production and distribution of video entertainment and advertising assets across major media platforms. We also provide definitions of common asset ID terminology.

This primer offers no endorsements of any methodology, nor is it comprehensive. Instead, it is meant to be a starting point in educating M&E industry participants about the many asset identification choices available, and the language used by the trade.

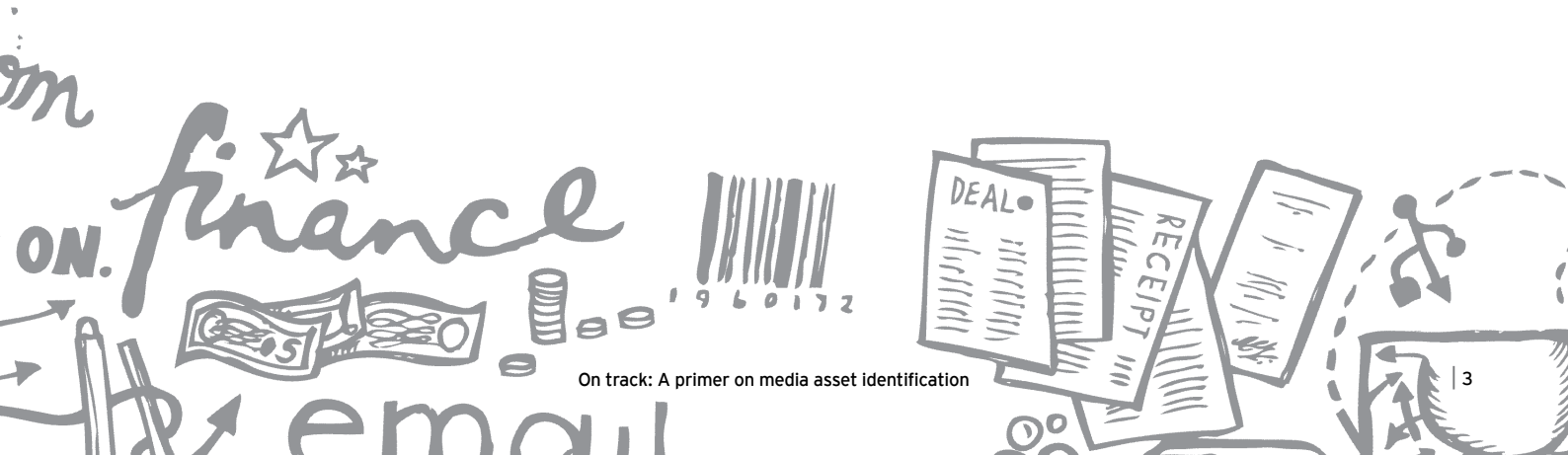



www
TECHNOLOGICAL
growth
watch
@the world
step one
step two
step three
Howdy
HOLA
HELLO

2.0 Asset identification schemata

2.1 Ad-ID

Who developed it	4A's and the Association of National Advertisers (ANA)
Who uses it	Advertisers and ad agencies.
What is it	A unique ID advertising asset coding system for use with all forms of advertising media (print, video, voice).
How is it generated	Codes are generated via a web-accessible database located at www.ad-id.org. Companies pay a fee per prefix and per item coded.
Metadata support	The Ad-ID registry database supports over 70 metadata fields including advertiser, product, brand, ad title, medium, agency and length/size.
Technical construct non-visual media (e.g., print, voice)	**ABCD** **Company prefix** Assigned to the registering company by Ad-ID; a company can have more than one prefix. **12345678** **8-digit unique code** Format of the unique code can differ depending upon advertiser preference; however, the last three digits of the code are always "overflow," which are used to prevent duplication of codes.
Technical construct visual media (e.g., film trailer, TV)	**ABCD** **Company prefix** Assigned to the registering company by Ad-ID; a company can have more than one prefix. **1234567** **7-digit unique code** Format of the unique code can differ depending on advertiser preference; however, the last three digits of the code are always "overflow." **H** **High Definition Indicator** Denotes an ad created in high-definition. If an ad is created in standard def, the H is eliminated and that digit is left blank.
Value chain	Advertisers register an advertisement with the Ad-ID system to receive a unique advertising identifier (the Ad-ID). The Ad-ID can be used throughout the advertising distribution and delivery value chain by the advertiser's digital asset management systems, media buying agency's sales systems, and broadcaster's traffic, scheduling and billing systems. As the advertisement moves through the value chain, the information associated with the ad (the metadata) is read and utilized by the various systems. Historically, the metadata of an ad was hand-typed into each system as it moved through the value chain. However, with the Ad-ID, the metadata can be downloaded from the registry, thereby improving the accuracy of the metadata as well as the accuracy of the ad's usage, including billing. Ad-ID leadership has indicated that approximately 21% (590) of parent companies that spend more than $5 million annually on advertising (2,855) are actively using Ad-ID.
Relevant references	▶ http://www.ad-id.org/help/help_detailNEW.cfm



2.2 Digital Object Identifier (DOI)

Who developed it	The International DOI Foundation (IDF)
Who uses it	Anyone who wants to create a reference to an object on a digital network.
What is it	A DOI is a name that references an object (e.g., book, photograph, audio recording, video recording, etc.) on a digital network (e.g., the internet). Similar to a URL, a DOI is a character string used to uniquely identify the object, but instead of pointing to a website, it points to where an object can be found.
How is it generated	Companies or individuals wanting to create a DOI would contact the DOI registration authority for the assignment of a unique registrant code (prefix). They would then register individual objects according to the construct below.
Metadata support	There is no metadata inherent in a DOI number; however the DOI registry maintains a data dictionary which can be used to assign metadata when registering the object and obtaining a DOI number.
Technical construct	10.1000 — **Prefix**: Assigned to the person or company registering the ID 123456xyz — **Suffix**: Unique number that is provided by the object's registrant, and is not determined by the IDF. Can be an ISAN, ISBN or some other standardized numbering scheme or a proprietary number. **Note:** ▶ A DOI name may be assigned to any item of intellectual property. ▶ There is no limitation on the length of a DOI name. ▶ The DOI numbering system requires the registrant to have its own numbering system (or use an existing system such as an ISBN) for the suffix. It does not generate a unique number on its own.
Value chain	Used primarily in academic publishing, a DOI is a way to establish a permanent citation for a piece of authored work (e.g., journal, article, book or thesis). Data and article citation services use a DOI as a means to point users to the work's location on the internet. DOI number schema is permanent and does not change even if the location of the object changes. It is notable that the DOI registry does not assign a unique number; DOI requires the registrant to have its own system in use.
Relevant references	▶ http://www.doi.org/about_the_doi.html

2.3 Entertainment Identifier Registry (EIDR)

Who developed it	Founded by MovieLabs, CableLabs, Comcast and Rovi
Who uses it	The registry has two categories of users: registrants who register media objects of different types and users who query the registry using EIDRs or other search criteria.
What is it	EIDR is a universal DOI that uniquely identifies an audiovisual object. It is similar to a UPC code that is used to identify physical packaged goods. EIDR can be used for both physical and digital video objects that are part of the movie and television supply chain.
How is it generated	The registry receives and processes registration requests from registrants. Users and applications can look up and search the registry. Registrants and look-up users can use the web interface or web services API to interact with the registry. A registrant submits objects for registration along with core metadata and information such as the type of object and relationship to other objects. EIDR uses a sophisticated de-duplication system to ensure that the object submitted to the Registry has not already been registered while allowing the registration of similar and related objects. If no duplicate object exists, the Registry generates an EIDR for the object and stores the new EIDR and the corresponding metadata in the Registry.
Metadata support	The metadata required and stored by the registry is restricted to those core elements that help uniquely identify the object that is being registered. EIDR does not provide metadata intended for consumers, extended or non-factual metadata (e.g., cast and crew, synopses, artwork, ratings), or a rights repository.
Technical construct	10.123 / 1234-5678-9ABC-DEFO - K Standard Prefix for EIDR Registry Unique Suffix for Each Asset Check Digit **Note:** EIDR is an opaque ID with all information about the registered asset stored in the central registry. EIDR is purely functional without any implication of ownership, making it persistent enough to remain the same despite any change in control or ownership of the underlying asset.
Value chain	EIDR provides the foundational namespace for all movie and television objects that are relevant to commerce. EIDR provides a registry to assign and store universally unique identifiers for a wide range of relevant movie and TV objects and enables programmatic interfaces to the registry for registering and searching records on a low-cost, non-profit basis. This supports creation of enhanced, value-added services including greater granularity in reporting down to the level of clips, composites and encodings, simplified universal search and discovery, and detailed consumption metrics for assets. EIDR is designed to be interoperable and work seamlessly in a complimentary manner with existing identifiers. EIDR is a universal DOI that uniquely identifies an audiovisual asset. It includes a de-duplication module to guarantee uniqueness and is interoperable with existing identification schemata. It also supports international content. However, commercial metadata, including metadata intended for consumers and extended non-factual metadata (e.g., cast and crew, artwork, ratings) are not tracked and EIDR also does not track rights as it is purely functional without any implication of ownership.
Relevant references	▸ http://eidr.org/

2.4 Industry Standard Commercial Identifier (ISCI)

Who developed it	4A's and the ANA
Who uses it	Advertisers and agencies implemented ISCI in 1969. ISCI is no longer supported but is still in use by many advertisers and agencies in place of "house" IDs.
What is it	ISCI is a manual advertising asset-coding system used by advertisers. It was formally withdrawn from the marketplace in October 2007 and replaced with Ad-ID.
How is it generated	ISCI previously licensed two, three or four alpha prefixes to advertisers, from which agencies created complete eight-digit alpha-numeric codes. ISCI maintained records of the prefixes and agencies tracked the complete codes.
Metadata support	The technical construct of an ISCI contains information about the advertiser and the spot as listed below.
Technical construct	ABCD └The first four characters are alphabetic, representing the advertiser (e.g., PEMX for Pepsi). 1234 └The next four characters are numeric and represent the spot. Different codes are used for different versions of the same spot (e.g., variations in language and length).
Value chain	Advertisers register their advertisements with the ISCI system to receive ISCI prefixes. They can then use the ISCI code throughout the value chain of their advertisements. The metadata of an advertisement is typed by hand into each system as it moves through the value chain. ISCI codes have been widely accepted by advertisers and broadcasters. The major drawback is that the IDs are not globally unique; people can tag commercials with their own ISCI code. This can result in collision if two parties decide to use the same code.
Relevant references	▶ http://www.teamservices.net/teamservices/files/u6/Ad-ID-Advertiser-Brief-080715.pdf

2.5 International Standard Audiovisual Number (ISAN)

Who developed it	The International Standards Organization (ISO)
Who uses it	Film, broadcast, cable and game producers, distributors and broadcasters.
What is it	A voluntary, unique numbering system and metadata schemata enabling the identification of any audiovisual work, including films, shorts, documentaries, television programs, sports events, advertising and also their related versions.
How is it generated	The ISAN Registration Authority (ISAN-IA) database automatically assigns a number when a registrant enters the information into the registry.
Metadata support	There is no metadata inherent in an ISAN number itself. The work that an ISAN references is often identified by a metadata set.
Technical construct	0000–3BAB–9352–0000–G–0000–0000–Q └Mandatory: 16 hexadecimal digits (0–9, A–F) followed by an alphanumeric check digit. └Optional: 8 hexadecimal digits Followed by a check character that can denote a specific version of the work (e.g., edits for length or content).
Value chain	The ISAN is used to identify a particular piece of work throughout the value chain — from concept, through production and distribution, to consumption. The use of the full 24-digit ISAN (16 mandatory digits and 8 optional) enables users to differentiate between various forms of a particular work (e.g., the digital version for use in digital movie theaters; the tape version used in analog theaters; the DVD sold for home entertainment; or the version delivered to cable companies for video on demand). The system also has been adopted worldwide. However, systems within the value chain may only be coded to utilize the 16 mandatory digits and not the full code denoting differing asset versions (V-ISAN). Further, some entities that have implemented ISAN have indicated that a perceived high cost and narrow scope (in terms of asset types covered), have limited widespread adoption in the media industry where multiple versions and a growing variety of asset types are the trend.
Relevant references	▶ http://www.isan.org/portal/page?_pageid=164,40165&_dad=portal&_schema=PORTAL

2.6 Unique Material Identifier (UMID)

Who developed it	Society of Motion Picture and Television Engineers (SMPTE)
Who uses it	Network and cable video producers, distributors and broadcasters.
What is it	Unique number used in production and post-production work to identify specific instances of media files. Can be used to identify either audio or video bit streams.
How is it generated	UMIDs do not require a central registration authority. They are generated by the studio/network that owns the asset, using a dictionary or registry of codes developed by SMPTE and ISO.
Metadata support	Metadata within the UMID includes fields such as type of format, length and country of user.
Technical construct	A series of 8 or 16 sets of two hexadecimal digits. Extended 64-byte UMID 06.0E.2B.34.01.01.01.01.07.02.00.00.00.00.00.00.00.00 **Basic 32-byte UMID** Contains the codes developed by SMPTE and ISO to denote type of media, length, and copy or version number **Optional 32-byte source pack** Contains the codes developed by SMPTE and ISO to denote the creation time and date, geospatial coordinates of the recording location and country/organization/user information
Value chain	A UMID is an internal number and is not likely to be used outside of the organization that generated it. A published work may well have both an ISAN and a UMID, but the UMID would not likely be widely distributed or publicized. The UMID numbering system has the ability to differentiate between different versions, or essences, of a media file. However, codes are not designed to be used outside of the production part of the value chain.
Relevant references	▶ http://www.digitalpreservationeurope.eu/publications/briefs/UMID_Unique%20Material%20Identifier.pdf

2.7 Universal Resource Locator (URL)

Who developed it	Internet Engineering Task Force (IETF)
Who uses it	Online publishers, users and advertisers.
What is it	A unique address for identifying and locating a resource (e.g., web page, image) on the internet.
How is it generated	The generic syntax provides a framework for new schemes for names to be resolved, using as yet undefined protocols.
Metadata support	There is no metadata information inherent in a URL.
Technical construct	resource_type://hostname.domain:port/filepathname#anchor └Resource or service type identifier (http, ftp) └Service hostname (computer system name) └Pathname of file to be fetched or program to be run └For html files, an optional anchor name within files where the display should start
Value chain	URLs enable locating a resource on a closed network or the internet in addition to identifying the resource. URL is the most widely used type of uniform resource identifier (URI). URLs can be highly customizable in terms of defining a path name. Content management systems can generate search engine-friendly URLs to assist users, and thereby reach an enhanced level of search engine optimization. However, the presence of duplicate URLs in the same document affects the crawling, indexing and relevance of search results.
Relevant references	▶ http://www.w3.org/Addressing/rfc1738.txt

Global Media & Entertainment Center

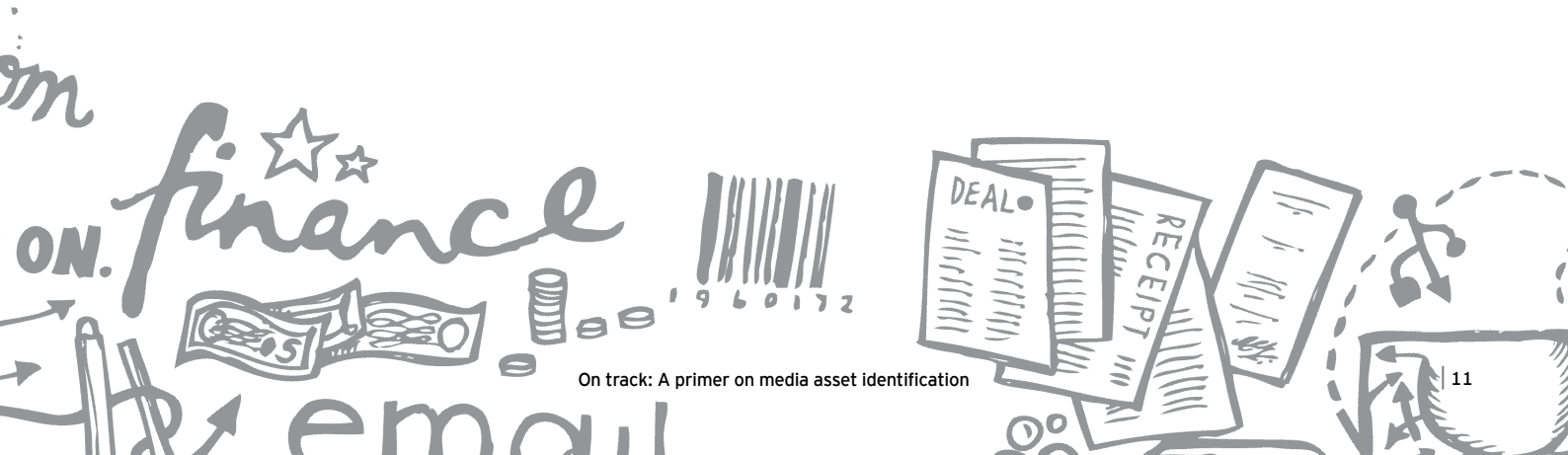
3.0 Engineering standards

3.1 Advanced Authoring Format (AAF)

Definition	AAF is an open-source file format which enables the exchange of data between multimedia authoring tools. The format allows interchange of essence data and metadata. Essence data includes picture, audio and video while metadata stores information on essence data. AAF was created by the AMWA. The format allows interoperability between different vendors and platforms for sharing multimedia information.
Implications	The AAF preserves metadata information while editing multimedia files during post production and authoring. This helps in versioning of files, improving production workflow, managing media, managing rights and changes to media and enhancing integration of creative tools with organizational systems. AAF's rich data model can also be applied beyond broadcast applications including the US Department of Defense for one of their surveillance applications.
Sources	▶ AMWA

3.2 Broadcast Exchange Format (BXF)

Definition	The BXF standard provides a single method of exchanging data among broadcast systems such as program management, traffic, automation, and content distribution. BXF standardizes the communication of three basic types of data exchange: ▶ Schedule and as-run information ▶ Content metadata ▶ Content movement instructions
Implications	BXF provides a standardized exchange of data among systems and thus, promotes integration among related systems. BXF also allows continuous, item-by-item, reconciliation between traffic and automation, removing tedious manual processes.
Sources	▶ SMPTE ▶ Broadcast Engineering



Global Media & Entertainment Center

4.0 Digital container formats

4.1 Extensible Metadata Platform (XMP)

Definition	XMP is an open standard, which was created in 2001 by Adobe Systems Inc. to store information about a file during the content creation process. Meaningful descriptions and titles, searchable keywords and up-to-date author and copyright information can be captured in a format that can be recognized by various software applications, hardware devices and file formats. XMP is serialized in Extensible Markup Language (XML) and stored by using a subset of the W3C Resource Description Framework (RDF). Thereby, customers can easily define their own custom properties and namespaces to embed arbitrary information into their files.
Implications	The XMP specification is adopted by various industry standard organizations including IPTC, DCMI, DISC and W3C. The most common metadata tags recorded in XMP data are those from the DCMI, which include the title, description, creator, etc. The standard is designed to be extensible, allowing users to add their own custom metadata to XMP data. XMP can be used in several file formats such as PDF, JPEG, JPEG 2000, GIF, PNG, HTML, TIFF, Adobe Illustrator, PSD, MP3, MP4, Audio Video Interleave, WAV and PostScript, Encapsulated PostScript, and is proposed for DjVu. In a typical edited JPEG file, XMP information is included alongside EXIF and IPTC Information Interchange Model data. XMP does not generally allow binary data types to be embedded. This means that any binary data one wants to carry in XMP, such as thumbnail images, must be encoded in an XML-friendly format. XMP metadata can describe a document as a whole (the "main" metadata), but can also describe parts of a document, such as pages or included images. Its architecture makes it possible to retain authorship and rights information (e.g., images included in a published document). It also permits documents created from several smaller documents to retain the original metadata associated with the parts.
Potential media applicability	▶ Television ▶ Mobile ▶ Broadband
Sources	▶ Adobe ▶ Metadata Working Group (MWG)

4.2 Material eXchange Format (MXF)

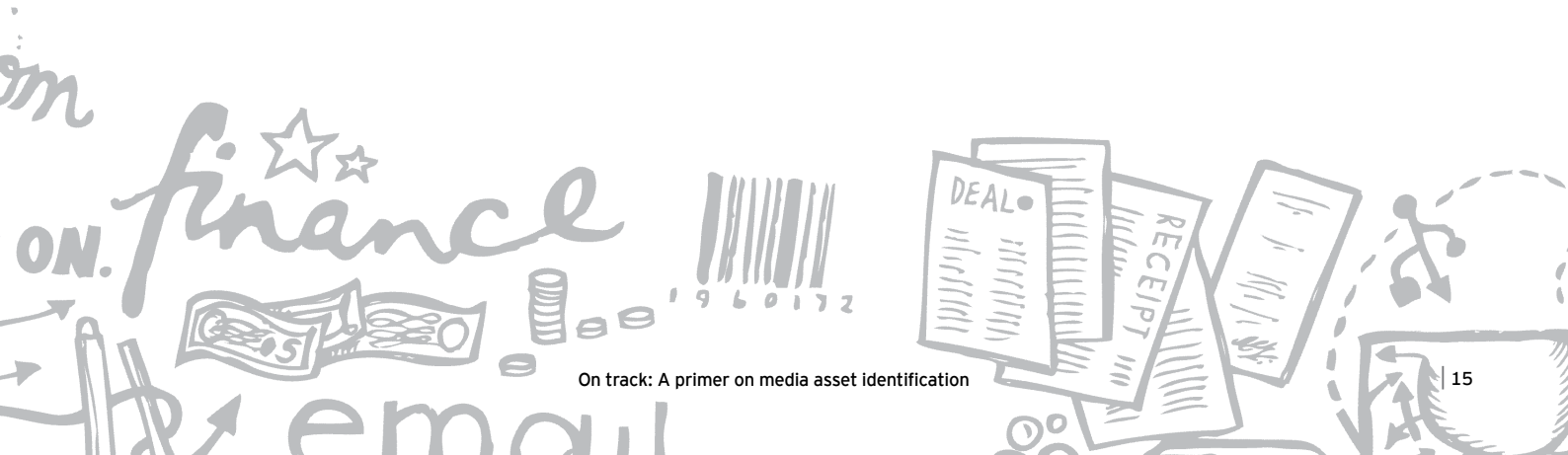
Definition	MXF is an open file format which acts as a "wrapper" of multimedia content and associated data and metadata.
Implications	The MXF format has been designed to have full metadata support and is applicable across multiple platforms for future professional video and audio applications. There have been some past interoperability problems with MXF as vendors implement fragmented parts of the standard. There are also some limitations to currently popular MXF export tools as they do not allow creation of a stereo AES (audio standard) file within the MXF wrapper or the addition of free text annotation to the MXF file created.
Potential media applicability	▶ Television ▶ Radio ▶ Broadband ▶ Print ▶ Mobile
Sources	▶ European Broadcasting Union (EBU) – Technical Review

Global Media & Entertainment Center

5.0 Key terms glossary

5.1 Application Programming Interface (API)

Definition	An API facilitates communication between different software programs by providing a standardized set of requests defined for various programs. An API is similar to the building blocks of a program. A programmer puts together the building blocks or a set of APIs to create a program. Most operating environments provide an API so that programmers can write applications that are consistent with the operating environment. APIs also help users by providing a consistent interface for various programs making it easier for users to learn new programs. An API includes a set of standard software interrupts, calls, functions and data formats that can be used by an application program to access network services, devices or operating systems. It may also refer to a complete interface, a single function or a customized set of routines developed by an organization. Thus, the scope of the meaning is usually determined by the context of the usage.
Implications	In the context of web development, API can be a defined set of Hypertext Transfer Protocol (HTTP) request messages. These messages are written usually in an XML or JavaScript Object Notation (JSON) format. Publishers have allowed web access to their APIs enabling web communities to create an open architecture for content sharing. The developer community can improve upon and collaborate with other communities to create more advanced applications or customize exiting ones to their requirements. Content that is created can be dynamically posted and updated in multiple locations on the web. Advantages of APIs: ▶ Developers can easily integrate remote tools and systems to accelerate their application development process. ▶ Companies do not have to pay for several different software applications as well as for the hardware to make them all work. ▶ The company that releases the API allows its customers to access their services in a new and more efficient manner. ▶ APIs can be protected from general use. For example, the API for Sony's PlayStation 2 was only available to licensed Sony developers. On the other hand, Microsoft's Windows API is freely available.
Sources	▶ Expedia Affiliate Network



5.2 Beacon

Definition	Web beacons are small strings of code that provide a method for delivering a graphic image on a web page or in an email message for the purpose of transferring data. A web beacon is also called a single-pixel, web bug, tracking bug, tracking pixel, pixel tag or clear gif image. Web beacons can recognize certain types of information on a computer such as the time and date of a page viewed, and a description of the page where the web beacon is placed.
Implications	A web beacon is downloaded whenever a user opens a graphical web page or an email. The browser requests the image to the server storing it, prompting the server regarding the download. As a result, web beacons can facilitate site traffic reporting, unique visitor counts, advertising auditing and reporting and personalization. Web beacons can also be used to deliver cookies or downloadable applications. In these situations, the code for the site being visited includes the same instruction to go to another server to fetch a small graphic file. However, instead of simply delivering the graphic file, the other server may also deliver a cookie or downloadable application. While web beacons are used in the same way in web pages or emails, they have different purposes: ▶ If the beacon is embedded in an email, a request is generated for the image when the user reads the email for the first time, and can also be requested each subsequent time the user loads the email. ▶ In the case of web pages, beacons may generate a "log file" record on the website's or third party's server. This may allow websites to better understand usage patterns and some limited characteristics about site visitors (e.g., the types of operating systems being used by visitors). As for all files transferred by using the HTTP, requests are made for web beacons by sending the server their URL, and perhaps the URL of the page containing them. Both the URLs contain information that can be useful for the server: ▶ The URL of the page containing the beacon allows the server to determine which particular web page the user has accessed. ▶ The URL of the beacon can be appended with an arbitrary string in various ways while still identifying the same object; this extra information can be used to identify the conditions under which the beacon has been loaded better. This extra information can be added while sending the page or by JavaScripts after the download. Web beacons can be used in combination with HTTP cookies like any other object transferred by using the HTTP.
Sources	▶ http://www.networkadvertising.org/networks/Web_Beacons_rev_11-1-04.pdf

5.3 Digital container format

Definition	A container or wrapper format is a meta-file format whose specification describes how different data elements and metadata coexist in a computer file. Containers are frequently used in multimedia applications.
Implications	Many multimedia data streams need to contain both audio and video data, and often some form of metadata that permits synchronization of the audio and video. Each of these three pieces of data may be handled by different programs, processes or hardware; but for the multimedia data stream to be useful in stored or transmitted form, they must be encapsulated together. A container is a way of "wrapping" audio, video streams with metadata into a single file.
Sources	▶ http://downloads.xiph.org/websites/xiph.org/container/

5.4 Digital watermarking

Definition	Watermarking is the process of embedding information into a digital signal in a way that is difficult to remove. The watermark "signal" may be audio, picture or video. If the signal is copied, then the information is also carried in the copy to enable the detection of copyright infringement. Also referred to as "encoding." Watermarking adds information, embedding it within a video and/or audio signal. A watermark is like a tattoo, permanently added to every frame of the digital media file. Watermarks may be visible (perceptible) such as network logos or station IDs used for branding, or invisible (imperceptible) for purposes of content protection where multiple copies of the same video could be uniquely identified within media asset management systems. Many identical pieces of a video can be created, each with a unique watermark, so that if one was uploaded without rights consent, the watermark would permit the media owner to identify the exact copy of the asset, and potentially, from where it was uploaded.
Implications	Watermarks are used for tracking individual assets, helping to identify content creator(s) and enabling rights management. Businesses find watermarking extremely useful because they are able to keep tabs on confidential recordings and videos, monitor distribution of sensitive material (e.g., such as previews of new movies for the Oscars), implement anti-piracy measures in digital cinema and drive content identification. Current digital watermarking methods embed codes so that the image can be altered without losing the ability to extract the watermark. Since digital watermarking is performed on uncompressed frames, it is typically carried out as part of a transcoding process. For a watermark to be useful there must be a way to extract it and compare it with known watermarks. Some vendors provide both watermarking and tracking services. ▶ **Audio watermarking**: Digital audio watermarking hides information in an audio file that is inaudible to the listener, and without affecting in any way the audio quality of the original file. The main use of an audio watermark is for protection of intellectual property rights, especially protection against online music piracy. Other uses of watermarking technology include embedding auxiliary information or metadata related to a particular song, such as lyrics, singer and other album information. One of the most secure techniques of audio watermarking is spread spectrum audio watermarking (SSW). New operating systems, equipped with the digital rights management (DRM) software, can extract the watermark from audio files before playing them on a system. The DRM software ensures that the user has paid for the song by comparing the watermark to the existing purchased licenses on the system. ▶ **Video watermarking**: A video watermark is an indelible pattern embedded in video content that is typically imperceptible to the eye. By embedding a unique watermark into video material content owners can identify copies of their materials. A video watermark can also be visible. Digital audio and video watermarking technologies find their application in acting as a deterrent against piracy, for forensic purposes and in broadcast distribution monitoring worldwide. They provide additional security for premium content delivered to PayTV and enhance the accuracy of audience measurement for radio, TV and catch-up TV.

5.4 Digital watermarking (continued)	
Implications (continued)	Watermarking is considered an active encoding technique in the industry since watermarks can be embedded into the assets prior to production and distribution. Below are cost and additional implications for the use of watermarks: ▶ **Cost**: Because a watermark system requires that the watermark embedding and reading be integrated into production or distribution systems, there is an up-front cost to adopting watermarking. However, once adopted and implemented, there is little ongoing cost, other than normal costs of maintaining the production and/or distribution systems associated with the watermarking embedding or reading process. ▶ **Efficiency**: Because a watermark contains metadata, different instances of the same original content may be easily differentiated. Metadata encoded into a watermark is called payload, and it can be both independent of or derived from the content. ▶ **Accuracy**: Watermarking systems are known to be highly accurate and the accuracy of watermarking systems is mathematically determined and demonstrated in practice. ▶ **Operations**: Digital watermarking ensures ownership is established up front and remains with the image even if it is manipulated, altered or distorted.
Sources	▶ http://www.digitalwatermarkingalliance.org/docs/papers/dwa_whitepaper_NewMobileApps.pdf ▶ http://www.licensestream.com/LicenseStreamPortal/Blog/post/2010/01/13/Embedded-Digital-Watermarking-vs-Digital-Fingerprinting.aspx

5.5 Fingerprinting

Definition	A technique in which software identifies, extracts and then compresses characteristic components of a video, enabling that video to be uniquely identified by its resultant "fingerprint." Also referred to as a "signature." Fingerprinting does not add any new information; rather it just analyzes the media asset to identify a unique match. All video fingerprints are stored in a reference database. Fingerprinting digital media works much like fingerprinting people. Any video clip can be compared to fingerprints on file to see if there is a match.
Implications	Applications of fingerprinting include broadcast and general media monitoring, copyright control, metadata tracking, behavioral modeling advertising, copy protection and forensics. A fingerprint is generated from a series of uncompressed frames and the fingerprint can incorporate metadata about the media along with the fingerprint pattern. Unlike watermarking, the fingerprint exporter does not generate a viewable file but rather a much smaller fingerprinting file documenting inherent characteristics of the media. A key implication for fingerprinting is that it is resolution- and format-independent and can be used to identify complete videos, portions of videos and short snippets of videos. Can also identify pieces of manipulated video content if included in a video mash-up. Fingerprinting is considered a passive technique in the industry. Below are cost and additional implications for the use of fingerprinting: ▶ **Cost**: Cost of a fingerprinting system is minimal up front but maintenance of the overall system is high since readers have to be enabled with updated algorithms related to coding and changes in the asset. ▶ **Efficiency**: The efficiency of a fingerprinting system is driven by the time it takes to search the reference database for a match. This time could potentially increase as the database gets larger in size with the growing volume of content. ▶ **Accuracy**: Fingerprinting systems have relatively small scale in existing applications so accuracy on a large scale is unknown. Fingerprinting accuracy is also impacted by the quality of sample images and the number of similar images or content that could be out there. ▶ **Operations**: A fingerprint requires a database for reference as the technology requires fingerprint to be sent to a server for comparison to the database. In isolation, the fingerprint contains no meaningful or actionable information; rather it provides information after resolution with the reference database. A fingerprint can be derived from content after it is distributed and the fingerprint content of an asset can be determined at any time during the life of the content including after production and distribution.
Sources	▶ http://www.digitalwatermarkingalliance.org/docs/papers/dwa_whitepaper_NewMobileApps.pdf ▶ http://www.licensestream.com/LicenseStreamPortal/Blog/post/2010/01/13/Embedded-Digital-Watermarking-vs-Digital-Fingerprinting.aspx

5.6 Identifier

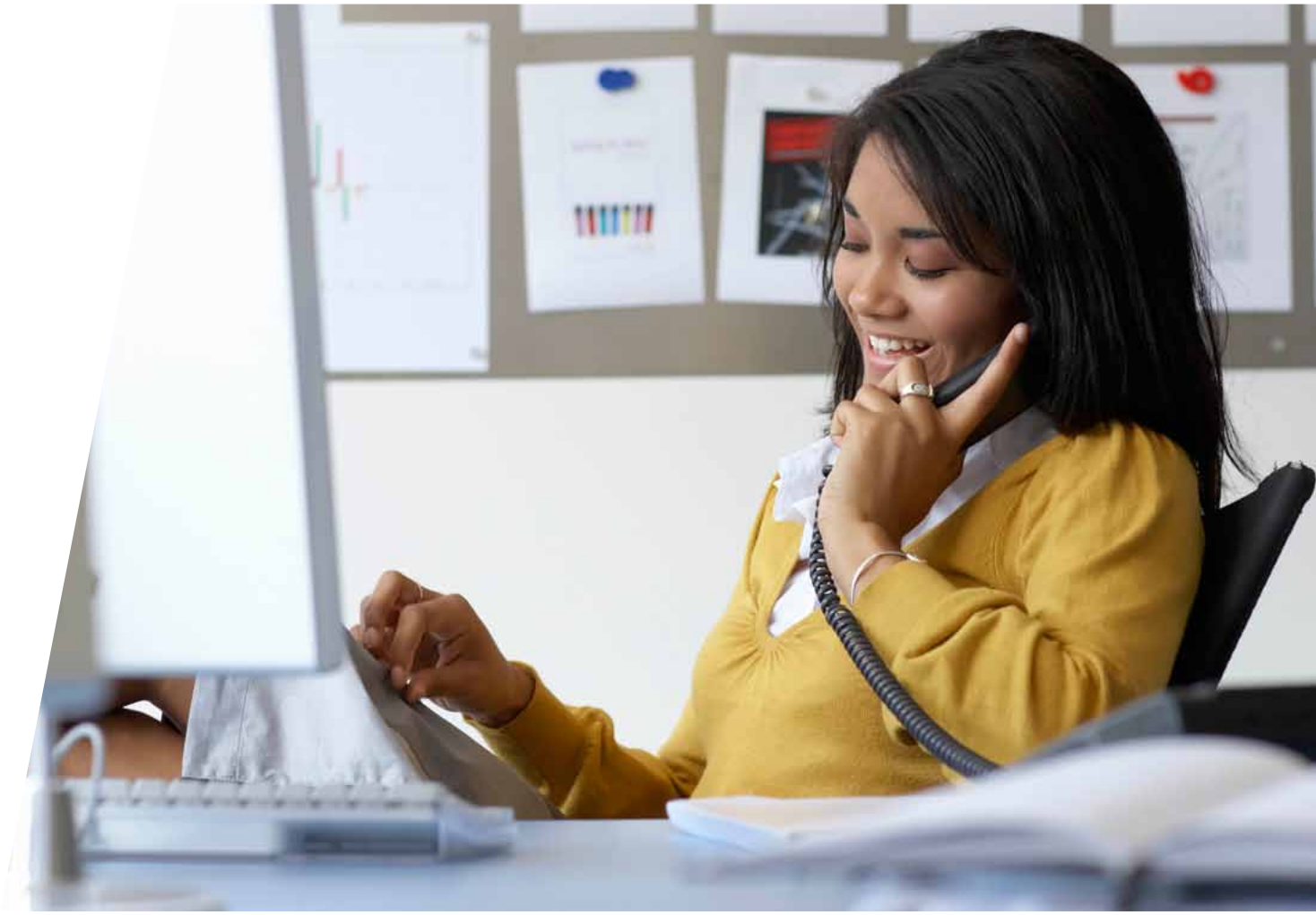
Definition	A unique expression in the form of a number, code or both to differentiate among a class of substances, items or objects.
Implications	Identifiers, including digital identifiers, provide a framework for persistent identification, managing intellectual content, managing metadata, facilitating electronic commerce and enabling automated management of media.
Sources	▶ http://www.doi.org/

5.7 Metadata

Definition	1. Bits and packets of data that can be used for a variety of purposes including market positioning, advertiser commercial tracking, viewership and referencing other data. 2. Data that is typically used to describe the property of a piece of content such as broadcast programming, VOD, interactive applications and advertising enhancements.
Implications	Metadata, in content identification, refers to data about the asset and is coded at various levels. Metadata elements grouped into sets designed for a specific purpose (domain or information resource) are referred to as metadata schemes. Different metadata schemes are developed as standards across disciplines. Relevant to the media industry, metadata standards have been developed as follows: ▶ **Digital Images**: NISO MIX technical metadata for digital still images is an XML schema for a set of technical data elements required to manage digital image collections. ▶ **Multimedia**: The Multimedia Content Description Interface MPEG-7 is an ISO/International Electrotechnical Commission (IEC) standard and specifies a set of descriptors to describe various types of multimedia information and was developed by the Moving Picture Experts Group. ▶ **Networked Resources**: (1) The Dublin Core Metadata Initiative has developed standards focused on networked resources. (2) DOI is a standard system for the identification and management of content on digital networks. The actual categories of metadata captured vary at two basic levels of "public" and "private" information. Data about the asset authorship and structure is often viewed as public data. However, data at the second level (e.g., cast crew, titles, and credits) might not always be public information. In the context of implementing a universal content identification standard, metadata categories need to be defined between public (what follows the asset) and private (what is retained within firewalls and authorized to be viewed by select users).
Sources	▶ http://www.cimm-us.org/ ▶ Fourthwall Media

5.8 Registry

Definition	A central location in an organization where information is stored and maintained in a controlled method.
Implications	A metadata registry typically has the following characteristics. ▶ Protected environment where only authorized individuals may make changes. ▶ Stores data elements that include both semantics and representations. ▶ Semantic areas of a metadata registry contain the meaning of a data element with precise definitions. ▶ Representational areas of a metadata registry define how the data is represented in a specific format, such as in a database or a structured file format (e.g., XML). The registry provides a means to standardize the information contained in it for use in other applications.
Sources	▶ http://www.isotc211.org/Registry_Reston/Registry_OBrien.ppt



5.9 Tag

Definition	A tag is a non-hierarchical keyword or term assigned to a piece of information such as an internet bookmark, digital image or computer file. This kind of metadata helps to describe an item and allows it to be found again by browsing or searching. Tags are generally chosen informally and personally by the creator of the item or by its viewer, depending on the system.
Implications	Tagging is carried out to perform functions such as aiding in the classification process, marking ownership, noting boundaries and indicating online identity. These may take the form of words, images or other identifying marks. Computer-based searching makes use of keywords as a rapid way of exploring records. Online and internet databases and early websites have traditionally deployed tags as a way for publishers to help users find related content. In 2003, the social bookmarking website De.licio.us provided a way for its users to add tags to their bookmarks and Flickr allowed its users a similar option to add tags to each of their pictures. The flexible and easy metadata made users' content and pictures highly searchable. Their success popularized the concept and other social software websites such as YouTube, Technorati and Gmail have implemented various forms of tagging. Users benefit from their own tagging which also enhances the experience of a website's greater community of users. In a traditional hierarchical system, the designer sets out a limited number of terms that are to be used for classification, and there is one correct way to classify each item. In a tagging system, there are an unlimited number of ways to classify an item. Instead of belonging to one category, an item may have several different tags. Some researchers and applications have experimented by combining a structured hierarchy and "flat" tagging to help in information retrieval. There are some special types of tags including the following: ▶ **Container tags**: Container tags are repeat element tags (e.g., a start and an end tag) that enclose other content. For example, a tag that indicates the enclosing content should be treated in a certain way, such as bolding, or italics, would have an opening tag and a closing tag to indicate when the treatment should start and end. By themselves, container tags cannot display or produce any output. They must enclose some content upon which they will act. ▶ **Triple tags**: Also known as a machine tag, this uses a special syntax to define extra semantic information about the tag, making it easier or more meaningful for interpretation by a computer program. Triple tags comprise three parts – a namespace, a predicate and a value. ▶ **Hash tags**: Short messages on services such as Twitter may be tagged by including one or more hash tags, which are words or phrases prefixed with a hash symbol (#), with multiple words concatenated. A user can search for the term # (word) and the tagged word will appear in the search engine results.

5.9 Tag (continued)	
Implications (continued)	There are some advantages and disadvantages of tags including the following: ▶ In a typical tagging system, there is no explicit information about the meaning or semantics of each tag, and users can apply new tags to an item as easily as when they apply older tags, and classify their collections of items in ways they find useful. Hierarchical classification systems can be slow to change, but a personalized variety of terms can present challenges when searching and browsing. ▶ When users can freely choose tags, the resulting metadata can include homonyms (the same tags used with different meanings) and synonyms (multiple tags for the same concept), which may lead to inappropriate connections between items and inefficient searches for information about a subject. Users can also choose tags that are different inflections of words (such as singular and plural), which can lead to navigation difficulties if the system does not include stemming of tags when searching or browsing. ▶ Tagging systems that are open to the public are also open to tag spam, e.g., when people apply an excessive number of tags or unrelated tags to an item (such as a YouTube video) to attract viewers. This abuse can be mitigated by using human or statistical identification of spam items. The number of tags allowed may also be limited to reduce spam.
Sources	▶ http://www.tourcms.com/support/webdesign/tags_container.php

5.10 Web crawler

Definition	A web crawler is a program or automated script which browses the internet in a methodical, automated manner. Other terms for web crawlers are crawlers, ants, automatic indexers, bots or web spiders. Processes with web crawlers are referred to as web crawling or spidering.
Implications	Many sites, search engines in particular, use web crawling as a means of providing up-to-date data. Web crawlers are mainly used to create a copy of all the visited pages for later processing by a search engine that will index the downloaded pages to provide fast searches. Crawlers can also be used for automating maintenance tasks on a website, such as checking links or validating HTML code. Furthermore, crawlers can be used to gather specific types of information from web pages, such as harvesting email addresses. A web crawler is one type of bot or software agent. In general, it starts with a list of URLs to visit, called seeds. As the crawler visits these URLs, it identifies all the hyperlinks in the page and adds them to the list of URLs to visit, called the crawl frontier. URLs from the frontier are recursively visited according to a set of policies. The characteristics of the web, such as large volume, fast rate of change and dynamic page generation, make crawling difficult. The large volume implies that the crawler can only download a fraction of the web pages within a given time, so it needs to prioritize downloads. The high rate of change implies that by the time the crawler is downloading the last pages from a site, pages have already been updated or even deleted. The number of possible crawlable URLs being generated by server-side software has also made it difficult for web crawlers to avoid retrieving duplicate content. Concepts related with crawling: ▶ **Focused crawlers**: The importance of a page for a crawler can also be expressed as a function of the similarity of a page to a given query. Web crawlers that attempt to download pages that are similar to each other are called focused or topical crawlers. The main problem in focused crawling is that in the context of a web crawler, we would like to be able to predict the similarity of the text of a given page to the query before actually downloading the page. ▶ **Restricting followed links**: A crawler may only want to seek out HTML pages and avoid all other types. Some crawlers may also avoid making requests for any resources that have a "?" in them (are dynamically produced) to avoid spider traps that may cause the crawler to download an infinite number of URLs from a web site. This strategy is unreliable if the site uses URL rewriting to simplify its URLs. ▶ **URL normalization**: Crawlers usually perform some type of modification and standardization of a URL in a consistent manner to avoid crawling the same resource more than once. This is also called URL canonicalization, such as conversion of URLs to lowercase, removal of "." and ".." segments and adding trailing slashes to the non-empty path component. ▶ **Path-ascending crawler**: Some crawlers are designed to download as many resources as possible from a particular website, so the path-ascending crawler was introduced to ascend to every path in each URL that it intends to crawl. Many path-ascending crawlers are also known as web-harvesting software, because they are used to "harvest" or collect all the content. ▶ **Politeness policy**: Crawlers can retrieve data much quicker and in greater depth than human searchers, so they can have a crippling impact on the performance of a site. For that reason ethical considerations must be taken into account when deciding where and how fast to crawl. ▶ **Parallelization policy**: A parallel crawler runs multiple processes in parallel. The goal is to maximize the download rate while minimizing the overhead from parallelization and avoiding repeated downloads of the same page by having a policy for assigning the new URLs discovered during the crawling process.

5.10 Web crawler (continued)	
Implications (continued)	▶ **Crawler identification**: Web crawlers typically identify themselves with a web server by using the user-agent field of an HTTP request. Website administrators examine their web servers' log and use the user agent field to determine which crawlers have visited the web server and how often. The user agent field may include a URL where the web site administrator may find more information about the crawler. It is important for web crawlers to identify themselves so that web site administrators can contact the owner if needed. ▶ **Crawling the Deep Web**: A large number of web pages lie in the deep or invisible web. These pages are typically only accessible by submitting queries to a database, and regular crawlers are unable to find these pages if there are no links that point to them. Google's Sitemap Protocol allows discovery of these deep-web resources. Deep web crawling also multiplies the number of web links to be crawled. Some crawlers only take some of the URLs.
Sources	▶ http://www.sciencedaily.com/articles/w/web_crawler.htm ▶ http://csjournals.com/IJCSC/PDF2-1/Article_49.pdf ▶ http://www.chato.cl/papers/crawling_thesis/effective_web_crawling.pdf

5.11 Wrapper

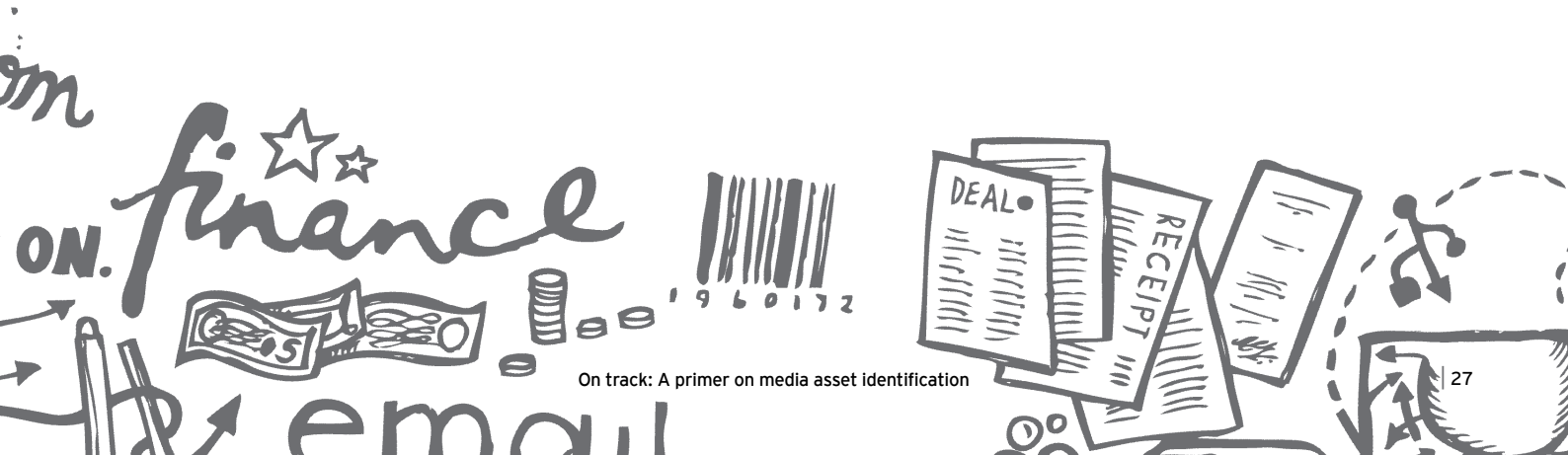
Definition	Wrappers are specialized program routines that automatically extract data from Internet websites and convert the information into a structured format.
Implications	A wrapper converts information that is implicitly stored on an HTML document into information that is explicitly stored as a data-structure for further processing. It can efficiently obtain the relevant information from an individual source by using a common query language. A wrapper provides a single uniform query interface to access a multiple information source. It performs a pattern-matching procedure, which depends on extraction rules. Creating a new wrapper for a new requirement varies in scale, depending on the text type, domain and scenario. Wrappers are used to extract online information from the source. The source can range from printed articles to the internet. The required information can be extracted from articles published in newspapers or by any print media by using conventional extraction methods. As the size and popularity of the internet is growing by the day, users find it difficult to extract the required information from web sites.
Sources	▶ http://www.knowlesys.com/articles/web-data-extraction/wrapper_definition.htm ▶ http://subs.emis.de/LNI/Proceedings/Proceedings29/GI-Proceedings.29-9.pdf

Global Media & Entertainment Center

Appendix A – Industry organizations

A.1 The Advanced Media Workflow Association (AMWA)

Background	The AMWA is an open, community-driven forum focused on the creation of timely, innovative, business-driven specifications and technologies for networked media workflows. The AMWA focuses on file-based workflows to benefit content creators including film, television, advertising, internet and post-production professionals. The AMWA works closely with standards bodies such as SMPTE. The AMWA continues its support of the Advanced Authoring Format (AAF) developed to enable content creators to easily exchange digital media and metadata across platforms. The AMWA strives for compatibility between AAF, BXF, MXF (AMWA is a co-creator) and XML. The AMWA operates multiple groups including the Media Services Architecture Group (MSAG), which is responsible for providing information on architectures and best practices for enterprise-level technology in digital media workflows. MSAG will fulfill its mission through the production of reference documents, educational information, examples and specifications where applicable.	
Key executives	Name	Role
	Brad Gilmer	Executive Director
Role in asset identification	The AMWA does not have a direct product offering for asset identification. However, it is active in driving projects focused on interoperability between technologies such as BXF files and MXF files.	
Relevant references	▶ http://www.aafassociation.org/index.shtml	



A.2 The Advanced Television Systems Committee (ATSC)

Background	ATSC is a non-profit industry group that has developed a set of standards for digital television transmission over terrestrial, cable and satellite networks. The ATSC standard was developed in the early 1990s by the Grand Alliance, a consortium of electronics and telecommunications companies that assembled to develop a specification for what is now known as high-definition television (HDTV). ATSC formats also include standard-definition formats, although initially only HDTV services were launched in the digital format.	
Key executives	Name	Role
	Mark Richer	President
	Jerry Whitaker	Vice President
	Lindsay Shelton Gross	Director of Communications
Role in asset identification	The HDTV standards defined by the ATSC produce wide screen 16:9 images up to 1920×1080 pixels in size, but many different image sizes are also supported. The reduced bandwidth requirements of lower-resolution images allow up to six standard-definition "subchannels" to be broadcast on a single 6 MHz TV channel. ATSC standards are marked A/x (x is the standard number). ATSC Standard A/53, which implemented the system developed by the Grand Alliance, was published in 1995; the standard was adopted by the Federal Communications Commission in the US in 1996. It was revised in 2009. ATSC Standard A/72 was approved in 2008 and introduced H.264/AVC video coding to the ATSC system. ATSC also incorporates 5.1-channel surround sound using the Dolby Digital AC-3 format. Numerous auxiliary datacasting services can also be provided. Many aspects of ATSC are patented, including elements of the MPEG video coding, the AC-3 audio coding and the 8VSB modulation. ATSC depends on numerous interwoven standards, e.g., the EIA-708 standard for digital closed captioning, which leads to variations in implementation.	

A.2 The Advanced Television Systems Committee (ATSC) (continued)	
Role in asset identification (continued)	There are various ATSC standards for various formats: ▶ **Audio**: Dolby Digital AC-3 is used as the audio codec, though it was officially standardized as A/52 by the ATSC. It allows the transport of up to five channels of sound with a sixth channel for low-frequency effects. ▶ **Video**: The ATSC system supports a number of different display resolutions, aspect ratios and frame rates. The formats are listed here by resolution, form of scanning and number of frames per second. For transport, ATSC uses the MPEG systems specification, known as an MPEG transport stream, to encapsulate data, subject to certain constraints. ▶ **Modulation and transmission**: ATSC signals are designed to use the same 6 MHz bandwidth as analog NTSC television channels. Once the digital video and audio signals have been compressed and multiplexed, the transport stream can be modulated in different ways depending on the method of transmission. ▶ **Mobile TV**: Mobile reception of digital stations using ATSC was difficult, if not impossible, until 2008. To overcome this, there were several proposed systems that reported improved mobile reception, which were submitted as candidates for a new ATSC standard, ATSC-M/H. After one year of standardization, the solution based on technology by LG Electronics was adopted. Like other worldwide open standards, the proposed ATSC mobile standards are backward-compatible with existing tuners, despite the fact that they were added to the standard after the original standard was in wide use. ▶ **Other systems**: ATSC coexists with the DVB-T standard and with ISDB-T. A similar standard called ADTB-T was developed for use as part of China's new DMB-T/H dual standard. While China has officially chosen a dual standard, there is no requirement that a receiver works with both the standards, and there is no support for the ADTB modulation from broadcasters or equipment and receiver manufacturers.
Recent news	ATSC replaced much of the analog National Television System Committee (NTSC) television system in the US on 12 June 2009 and will replace NTSC by 31 August 2011 in Canada, 31 December 2015 in Mexico and 1 January 2019 in El Salvador.
Relevant references	▶ ATSC website

A.3 American Association of Advertising Agencies (4A's)

Background	4A's is the national trade association of the advertising agency business in US.	
Key executives	Name	Role
	Nancy Hill	President and CEO
	Laura Bartlett	CFO and COO
	Sharon Napier	Secretary-Treasurer, BOD
	Chuck Porter	Chairman, BOD
Role in asset identification	The 4A's is one of two major developers and sponsors (along with the ANA) of the Ad-ID advertising unique ID coding system.	
Relevant references	▶ http://www.aaaa.org/Pages/default.aspx ▶ http://www.betteradvertising.com/daa_release.html	

A.4 Association of National Advertisers (ANA)

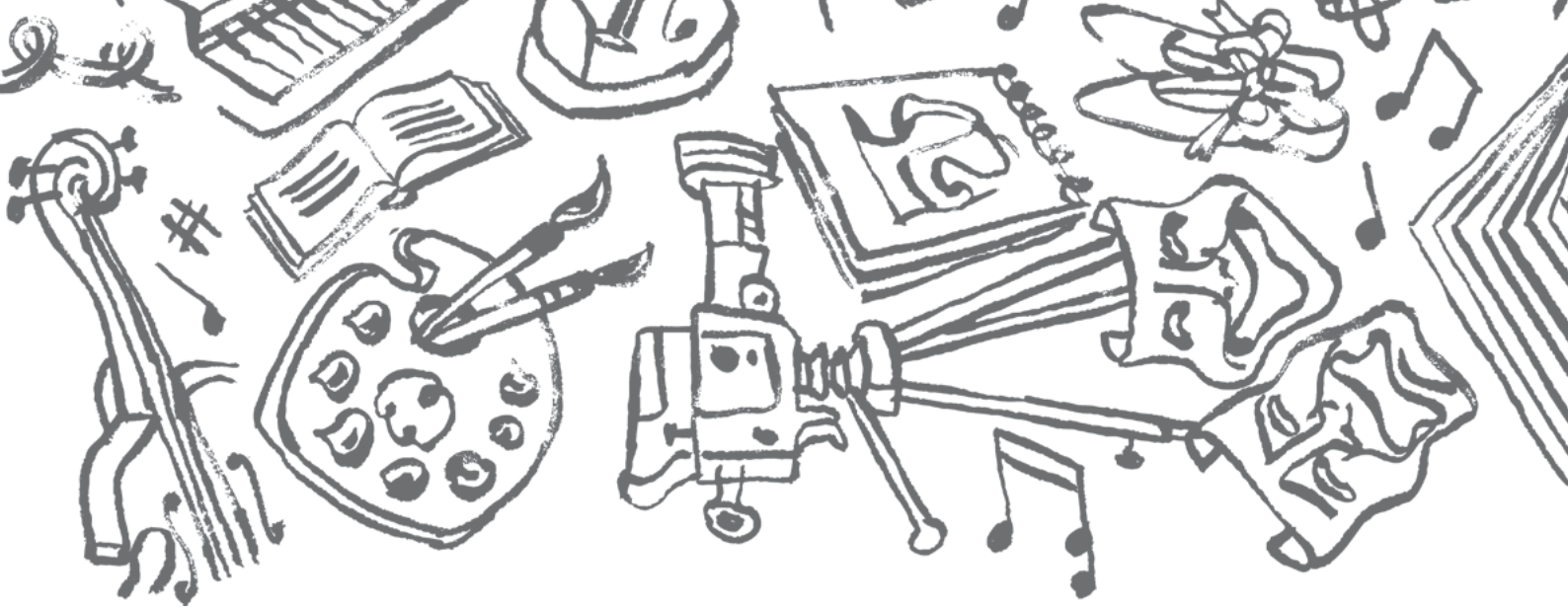
Background	ANA is the advertising industry's oldest trade association. Founded in 1910 to safeguard and advance the interests of advertisers and consumers, the ANA leads the marketing community by providing its members insights, collaboration and advocacy. Its membership includes 400 companies with over 9,000 brands.	
Key executives	Name	Role
	Bob Liodice	President and CEO
	Christine Manna	COO
	Duke Fanelli	SVP, Marketing & Communications
	Barry Garbarino	Director of Marketing
	Christine Manna	CFO and COO
	Robert Rothe	SVP and CIO
	Kristina Sweet	Senior Director, Sponsorship and Media Sales
	Kristen McDonough	Senior Director, Conferences and Forums
Role in asset identification	The ANA is one of two major developers and sponsors (along with the 4A's) of the Ad-ID advertising unique ID coding system.	
Recent news	**February 2010**: The ANA and The Nielson Company announced potential path for providing individual commercial ratings.	
Relevant references	▶ http://www.ana.net/ ▶ http://www.ana.net/content/show/id/575	

A.5 CableLabs

Background	CableLabs was founded in 1988 by cable operating companies and is a non-profit research and development consortium dedicated to pursuing new cable telecommunications technologies and to helping its cable operator members integrate those technical advancements into their business objectives. CableLabs works with members to determine what service requirements are to be supported by new technologies and new services. CableLabs is also involved in a few key technology projects including: ▶ **OpenCable**: OpenCable represents an effort to create a common platform for interactive services, programming, and advertising on retail and cable devices. OpenCable specifications describe an interactive digital cable platform comprised of baseline core functional requirements for digital cable ready "host" devices, a middleware comprising a set of common APIs, hardware interfaces between host devices and a removable CableCard, copy protection and security requirements, and optional extensions for host devices, including Home Networking, and DVR. ▶ **DOCSIS**: Cable modems based on Data Over Cable Service Interface. DOCSIS is the most successful and cost effective way to provide high speed data services is via cable modems compliant with the DOCSIS specifications. ▶ **VOD Metadata 3.0**: The VOD Metadata project is a television industry effort to specify the metadata and interfaces for distribution of Video on Demand (VOD) material from multiple content providers to cable operators. The project benefits cable operators and ultimately consumers by lowering encoding costs incurred by content providers, creating interoperability between different vendors' VOD systems. The VOD Metadata 3.0 specification is a CableLabs specification for descriptive data associated with a package of VOD content. The metadata is used in MSO and programmer VOD systems today, but in the future will assist in the delivery of prospective ad products for the VOD space, or in adding greater addressability to different types of ads.
Key executives	Name: Role Brian L. Roberts: Chairman Chairman and CEO, Comcast Corporation Glenn A. Britt: Vice Chairman of Cablelabs, President and CEO of Time Warner Cable Thomas M. Rutledge: Treasurer Chief Operating Officer, Cablevision Systems Corporation
Role in asset identification	CableLabs has not developed specific asset identification products. Instead, it is one of the founding members of a new international coalition that has formed the Entertainment Identifier Registry (EIDR).
Relevant references	▶ http://www.cablelabs.com/

A.6 Interactive Advertising Bureau (IAB)

Background	The IAB is an advertising business organization that comprises around 460 media and technology companies in the US. The organization works toward the growth of the interactive advertising marketplace, its share of the total marketing spend and that of its members' total sharing spend. Working with its member companies, the IAB evaluates and recommends standards and practices, and conducts research on interactive advertising.	
Key executives	Name	Role
	Randall Rothenberg	President and CEO
	Bob Carrigan	Chairman
	Peter Naylor	Vice Chairman
	Bruce Gordon	Treasurer
	Joseph Rosenbaum	Secretary
	Patrick Dolan	Executive Vice President and COO
	David Doty	Senior Vice President and CMO
	Sherrill Mane	Senior Vice President, Industry Services
	Mark Goldman	Senior Director, Finance and Administration
Role in asset identification	The IAB, in conjunction with its member companies, evaluates and recommends standards and practices for interactive advertising, including measurement.	
Recent news	**October 2010**: Five advertising trade groups, including the IAB, launched a new program encouraging members to use an advertising option icon, which is connected to an opt-out mechanism for targeted advertising, alongside online advertisements. The program provides consumers with enhanced control over the collection and use of data related to their web viewing for behavioral advertising purposes.	



A.6 Interactive Advertising Bureau (IAB) (continued)	
Recent news (continued)	**October 2010**: The Coalition for Innovative Media Measurement (CIMM) initiated a study to examine the feasibility of an open standard for tracking and identifying entertainment content and advertisements across traditional and digital media platforms. It is launching its "Project: TAXI," which stands for "Trackable Asset Cross-Platform Identifier," to respond to the growing difficulty that content rights holders and advertisers face in maximizing the value of their assets. CIMM has partnered with the IAB, ANA and 4A's on Project: TAXI. **April 2010**: IAB, in conjunction with the advertising network trade group Network Advertising Initiative, launched a new initiative, Control Links for Education and Advertising Responsibly (CLEAR), aimed at increasing consumer awareness relative to behavioral-targeting advertising practices. Through this initiative, CLEAR informs people about the reason they are receiving a particular web advertisement. It also advocates that publishers and advertising networks run notices alongside banner advertisements enabling users to access additional information about the origin of an advertisement and allowing users to opt out, if they desire to do so.
Relevant references	▶ http://www.computerworld.com/s/article/9189338/Online_groups_introduce_labeling_for_targeted_ads?taxonomyId=16 ▶ http://www.broadcastingcable.com/article/458703-CIMM_Seeks_Open_Format_to_Track_Content.php?rssid=20102 ▶ http://eon.businesswire.com/news/eon/20101213006894/en/Adobe/Omniture/Search%26Promote

A.7 Media Rating Council (MRC)

Background	The MRC is a non-profit industry group that audits media research companies to verify the accuracy of their audience research. Its industry group members include television and radio broadcasters, print organizations, advertisers, internet organizations, advertising agencies and industry trade associations.	
Key executives	Name	Role
	George Ivie	Executive Director and CEO
	Billy McDowell	Chairman, BOD
Role in asset identification	The MRC audits audience measurement companies to determine if their practices conform to accepted standards. The MRC engages public accounting firms, such as Ernst & Young, to perform these audits on its behalf.	
Relevant references	▶ http://mediaratingcouncil.org/	

A.8 MovieLabs

Background	Motion Picture Laboratories, Inc. (MovieLabs) is a non-profit research and development joint venture launched by six major motion picture studios: Paramount Pictures Corporation, Sony Pictures Entertainment Inc., Twentieth Century Fox Film Corporation, Universal City Studios, Walt Disney Pictures and Television and Warner Bros. Entertainment, Inc. It provides universities, corporations, startups and network service operators with technical guidance and funding to enable them to explore innovative technologies in the distribution and use of motion pictures as consumer media. MovieLabs is funded by a multi-year commitment from its members.	
Key executives	Name	Role
	Steve Weinstein	President and CEO
	James Helman	Chief Technology Officer
	Raymond Drewry	VP EMEA Operations, Principal Scientist
	M. Kip Welch	Vice President, Business Development
	Craig Seidel	Vice President, Digital Distribution
Role in asset identification	MovieLabs is co-developing and deploying the Entertainment Identifier Registry (EIDR) for digital content.	
Recent news	**October 2010**: Movielabs, CableLabs, Comcast and Rovi Corporation partner to create the EIDR, a non-profit company that provides a uniform approach to catalog entertainment assets with unique identifiers. Built on the established DOI system, EIDR will serve as an open, standards based identification system enabling improved rights tracking and reporting. The EIDR identifier functionality is similar to that of product UPC codes and for book ISBN codes. It will help increase workflow efficiencies and lower the risks of misidentification. EIDR is working closely with the Center for National Resource Initiatives (CNRI) as its software integrator and professional services firm as CNRI has supported academic journals and the military with similar systems in the past.	
Relevant references	▶ http://www.movielabs.org/ ▶ http://eidr.org/ ▶ http://www.homemediamagazine.com/electronic-delivery/industry-group-launches-entertainment-registry-20970 ▶ http://blogs.gartner.com/mike_mcguire/2010/11/07/eidr-a-much-needed-standard-for-licensing-and-online-distribution/	

A.9 National Association of Broadcasters (NAB)

Background	The NAB is a trade association for broadcasters. The NAB delivers value to members through advocacy, education and innovation. ▶ **Advocacy**: The NAB is chief advocate of broadcasters and ensures that policymakers are informed on the issues that impact the broadcasting industry. ▶ **Education**: The NAB provides free public service materials to help broadcasters grow in their careers, promote diversity and strengthen their business. ▶ **Innovation**: The NAB is helping broadcasters find innovative ways to deliver high-quality content and services. In addition, the NAB is a founding member of the ATSC.
Key executives	Name: Role Gordon Smith: NAB President and CEO Steven W. Newberry: BOD President and CEO, Commonwealth Broadcasting Corporation Jack Sander: BOD Senior Advisor, Belo Corporation Jack Abernethy: BOD CEO, Fox Television Stations, Inc.
Role in asset identification	The NAB does not have a direct product offering for content identification. Instead, it is one of the founding members of ATSC which sets the standards for digital television and related content identification.
Recent news	**September 2008**: The ATSC recommended ISAN as a content identification standard. ATSC aims to bring the method for content identification in line with those defined in MPEG-2 systems. ATSC is the Standard for Digital Broadcasting, replacing the analog NTSC system, adopted in USA, Canada and in many countries in Latin America such as Mexico, Honduras and Puerto Rico. Open Mobile Video Coalition (OMVC), an alliance of U.S. commercial and public broadcasters, formed to accelerate the development and rollout of Mobile Digital Television (DTV) products and services. The OMVC is committed to maximizing and developing the full potential of digital television spectrum.
Relevant references	▶ http://www.nab.org/about/default.asp ▶ http://www.isan.org/docs/newsletter_october_2008.pdf

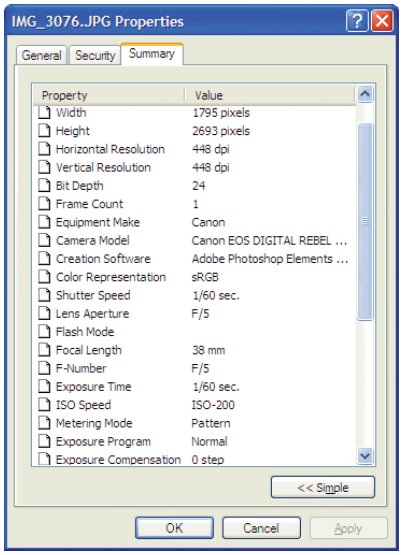
A.10 The Society of Motion Picture and Television Engineers (SMPTE)

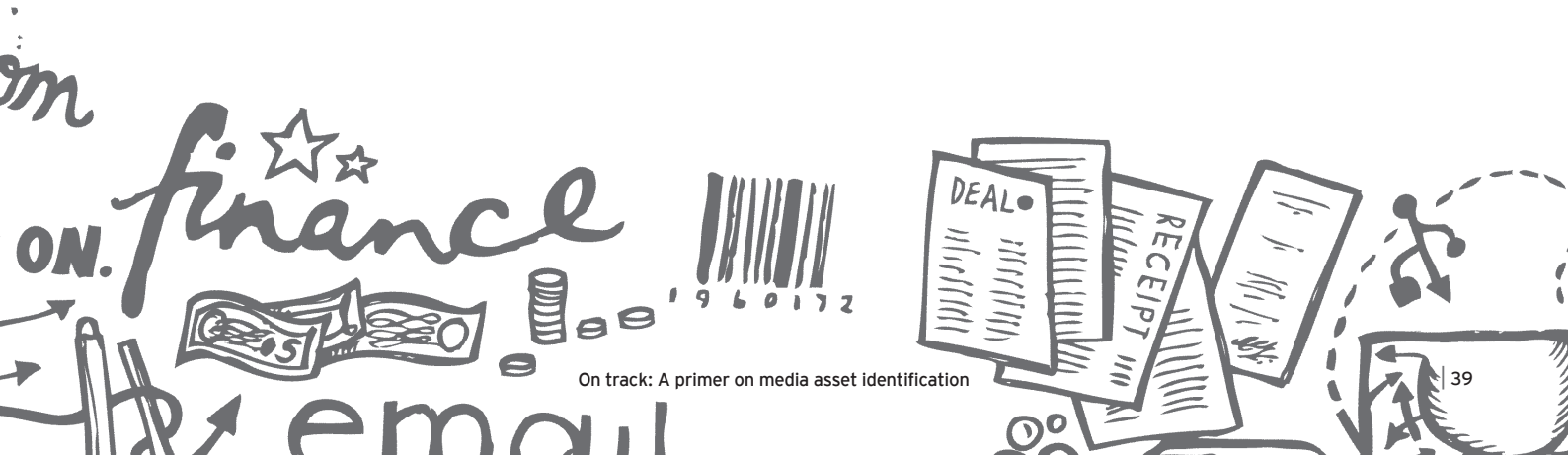
Background	SMPTE is the leading technical society for the motion imaging industry. It was founded in 1916 to advance theory and development in the motion imaging field. Today, SMPTE publishes ANSI-approved standards, recommended practices, and engineering guidelines, along with the highly regarded *SMPTE Motion Imaging Journal* and its peer-reviewed technical papers. SMPTE strives towards its goal through: ▶ Developing industry standards. ▶ Enhancing education through seminars, exhibitions, and conferences. ▶ Communicating the latest developments in technology. ▶ Promoting networking and interaction. The SMPTE standards cover a broad context including content creation, distribution, archive and playback, with participants that include equipment manufacturers, software developers, studios, European broadcasters and other content creators. The SMPTE has formed various committees to drive and publish standards across the industry including digital cinema, which includes the metadata and registers committee and regulates the metadata format. The SMPTE structure includes projects and initiatives to establish industry standards and technology advancements through: ▶ **SMPTE Standards**: The SMPTE has over 400 published standards. The most advocated-for standards include all film and television transmission formats, including digital, physical interfaces for transmission of television signals and related data; the SMPTE color bar test pattern and other diagnostic tools; and MXF. ▶ **Technology Committees**: The scope of these committees is to develop SMPTE engineering documents; review existing documents to ensure that they are current per established engineering practices and compatible with other international engineering documents, where possible; recommend and develop test specifications, methods and materials; and prepare tutorial material on engineering subjects for publication in the *SMPTE Motion Imaging Journal* or for other means of dissemination benefiting the Society and the industry.
Key executives	Name: Role Peter Lude: President Wendy Aylsworth: Executive Vice President Hans Hoffman: Engineering Vice President Robert P. Seidel: Financial Vice President
Role in asset identification	SMPTE does not have a direct product offering for content identification. However, it is heavily involved in publishing industry standards relevant to content identification through its Metadata Registries technology infrastructure committee.
Relevant references	▶ http://www.smpte.org/standards/committees/ ▶ http://www.oscars.org/science-technology/council/projects/metadata-symposium/media/dmpms_08_broome.pdf

Global Media & Entertainment Center

Appendix B – Selected case studies

B.1 Exchangeable image file format (Exif)

Introduction	Exif is metadata embedded in an image by a digital camera when the image is first captured. The format is created by Japan Electronics Industries Development Association (JEIDA) to encourage interoperability between imaging devices. Most Exif fields are write-protected and cannot be edited by software applications; this helps secure the integrity of the original photo capture information. Exif metadata is supported in formats such as JPEG, TIFF, Rev. 6.0 and RIFF WAV. Exif data includes the following information: ▶ Date and time settings for the image. ▶ Camera setting information which includes static information such as the camera model and make and dynamic information such as orientation/rotation, aperture, shutter speed, focal length, metering mode and ISO speed information. ▶ A thumbnail for previewing the picture on the camera's LCD screen, in file managers, or in photo manipulation software. ▶ Descriptions and copyright information.
Specification	The Exif image file specification stipulates the method of recording image data in files, and specifies the following items: ▶ Structure of image data files. ▶ Tags used by this standard. ▶ Definition and management of format versions. **Features of the Exif image file specification include:** 



B.1 Exchangeable image file format (Exif) (continued)	
History and origin	Before the digital era, the photographer had to manually record the settings of each shot to understand what combination of settings produced what effects. Fujifilm originally proposed the Exif image file format in 1994 to standardize the capture of information relating to the images. JEIDA announced Exif version 2.1 in June 1998 which was implemented as a format for storing metadata on digital images in June 1998. It has been embraced by a number of other digital camera manufacturers and though the specification is not currently maintained by any industry or standards organization, its use by camera manufacturers is nearly universal and it is under consideration as an ISO standard.
Competing standards	A new standard for digital cameras – Camera Image File Format (CIFF) – was also proposed in June 1998 for storing digital still images. Twenty camera manufacturers led by Canon and including Nikon, Olympus and Pentax supported this format stating that it is more compact than the Exif format. The CIFF standard was incompatible with Exif which was backed by major film manufacturers Kodak and Fuji. In October 1998, JEIDA proposed to merge the two formats into a new format called Digital Camera File format (DCF). It was intended to tap the best of each format and define file-handling specifications to assure compatibility at the hardware level. The final format after the unification still faced competition from proprietary formats such as those from Sony, Casio and FlashPix, which were promoted by Microsoft, Live Picture and Hewlett-Packard. The Exif supporter formed a group called the Exif Supporters Group (SEG). SEG consisted of 11 Japanese companies: Casio, Fuji Photo Film, Konica, Minolta, Nikon, Olympus, Ricoh, Sanyo, Seiko Epson, Sharp and Toshiba. Except for Fuji, Ricoh, Sharp and Toshiba, the other members also belonged to the CIFF camp. SEG entered into an agreement on the optional specifications that were left to each manufacturer when Exif Version 2 was adopted as a standard by the JEIDA in November 1997, so that digital cameras will have enhanced media compatibility. SEG standardized 15 items, including thumbnail and data-storage formats.

B.1 Exchangeable image file format (Exif) (continued)			
Timeline for different versions of Exif format	**Version**	**Date**	**Comment**
	1.0	October 1995	First edition established image data format definitions, structure and attribute (tag) information and tag definitions
	1.1	May 1997	Added tags and operating specifications
	2.0	November 1997	Added sRGB (standard for monitors, printers and the internet) color space, GPS, compressed thumbnails and audio files
	2.1	December 1998	Added DCF (design rule for camera file system) interoperability tags
	2.2	April 2002	Developed tags for improved print quality (contrast, sharpness, etc.) and added more tags pertaining to positioning and GPS
	2.21	September 2003	Added and corrected Exif 2.2 content in line with revision of DCF 2.0, added operational guidelines, corrected notations on image data pixel composition and pixel sampling, corrected misprints and omissions throughout the text.
	Unified v2.21	September 2009	Merged and added a portion of Exif 2.21 to Exif 2.2, added guidelines for handling Exif/DCF issued by CIPA
	2.3	April 2010	Added and revised tags; restructures main standard text, guidelines, explications, etc. of Exif Unified Version 2.21; clarified specification levels and revised the scope of application, supplemented explanations and adjusted formats for entire text.
Conclusion	Despite the existence of multiple file formats for storing image file specifications, Exif is widely used due to its various benefits such as standardized color description and compatibility with high-end color management systems. However, the Exif standard is not an officially accepted standard for storing image metadata. The standard suffers from various drawbacks such as its legacy file structure, limited supporting file formats, limits on color depth, no provision for video file information and absence of time-zone information. Nevertheless, Exif remains the most popular standard for image file data capture.		

B.2 Universal Product Code (UPC)

Introduction	UPC is a type of barcode that is widely used in US and Canada for tracking trade items in stores. The barcode was initially developed for grocery stores to speed up the checkout process, reduce errors and improve inventory management. The adoption spread to other industries due to its huge savings potential. UPC was introduced by the Uniform Code Council, Inc., an organization that until 1972 was known as the Uniform Grocery Product Code Council. The Uniform Code Council then merged with European Article Numbering (EAN) International which was further merged into Global Standards One (GS1). GS1, thereby, became the single worldwide origination point for UPC and EAN numbers.
Specification	A UPC symbol has two parts: 1. A machine-readable bar code 2. Human-readable 12-digit code  **First six digits:** Manufacturer identification number provided by GS1. **Next five digits:** Item number (varies by product, size, and color etc.) **Last digit is a check digit calculated as follows:** **Step 1**: Add the value of all digits in odd positions and multiply it by 3. [(6+9+8+0+0+9=32)*3=96] **Step 2**: Add the value of all digits in even positions and add the resulting number to the number obtained in Step 1. [(3+3+2+0+3=11)+96 =107] **Step 3**: To obtain the check digit, find the number to be added to reach the closest multiple of 10 for the number in Step 2. This number is the check digit. (107+3 = 110) In general, every item the manufacturer sells, as well as every size package and every repackaging of the item, needs a different item code. A person employed by the manufacturer, called the UPC coordinator, is responsible for assigning item numbers to products, making sure the same code is not used on more than one product, retiring codes as products are removed from the product line, etc. The last digit of the UPC code is called a check digit. This digit lets the scanner determine if it scanned the number correctly or not.

B.2 Universal Product Code (UPC) (continued)	
History and origin	The origin of bar code dates back to 1949 when a US engineer, Joe Woodland, decided to automate the supermarket-checkout process since the manual system was slow, error-prone and costly. In 1952, Joe, along with a partner, patented their system. However, the system required wide collaboration by the industry players: manufacturers, retailers, consumers and labor unions. Throughout the 1960s, the development of a universally acceptable bar coding system remained stagnant due to uncoordinated efforts from various participants. Mostly, disagreements originated between manufacturers and retailers on the size of the code and its compatibility with existing product codes. Manufacturers wanted a longer code to capture greater information while retailers focused on shorter codes as they might have to key in the information at checkout. In August 1970, representatives of member organizations of the Distributor and Manufacturer Association formed The Grocery Industry Ad Hoc Committee, which was charged with studying and reporting the economic potential of a UPC and identifying potential roadblocks to implementation. The Ad Hoc Committee was comprised of 18 respected executives from companies representing all areas of the grocery supply chain. In 1972, the Symbol Standardization Committee was formed to evaluate alternative symbol proposals. The evaluation process included presentations to various manufacturing groups and key executives. The process continued through 1972 and early 1973 and eventually the committee approved the symbol, in its current form, proposed by IBM.

B.2 Universal Product Code (UPC) (continued)	
Roadblocks to adoption	Convincing a gamut of retailers, manufacturers, wholesalers, consumers and legal organizations to adopt a single code to identify products was a complex task entailing the coordination among all the groups. Following were the initial concerns emerging from various entities: ▶ **Economic potential**: Various parties expressed concerns over the cost of implementation of UPCs compared with the potential cost savings. The parties doubted whether the benefits could be more than the cost of implementation. In addition, the success of a UPC system rested on the acceptance by a critical mass of retailers/manufacturers leading to a "chicken and egg problem." All the parties depended on each other to move forward with the adoption of UPC for realizing maximum benefit from the technology. ▶ **Rapid technology changes**: Technology, which changed rapidly during that time, led to further concerns about whether it was too early to establish a symbol standard. The threat of emergence of a technologically-superior standard, making the existing one obsolete, put a number of stakeholders off from adopting UPC. ▶ **Opposition from trade unions**: The implementation of UPCs was vehemently opposed by the trade unions, which feared huge job losses at the retailer end due to implementation of the technology. ▶ **Consumer skepticism**: Consumer advocacy groups feared that the absence of individual price markings on the products would encourage misinformation by retailers about a product's price. The concerns led to the enforcement of various price marking legislations, thereby, lengthening the payback period for the investment in the UPC system.

B.2 Universal Product Code (UPC) (continued)	
Success factors	Following were the critical success factors for implementation of the UPC system amid a number to similar failed attempts: ▶ **Focusing on real benefits**: The Ad Hoc Committee focused on real opportunities which a UPC system was likely to bring, instead of focusing on the futuristic possibilities. The committee highlighted the immediate benefits, such as potential savings in checker productivity, automated ordering, reduced shrinkage and improved control systems. Improved savings in the short term helped demonstrate the value of UPC to late adopters. The committee conducted time studies and work standards to accurately establish the quantum of savings for implementing UPC. ▶ **Conservative assessment of intangible benefits to minimize opposition**: Realization of a few intangible benefits such as savings from inventory management depended upon altered structure of supplier relationships. These benefits could have attracted opponents who might have repelled UPC's implementation by terming the proposal as speculative. Thus, the Ad Hoc Committee decided to account for only 25% of the estimated intangible benefits likely to be realized from UPC. ▶ **Addressing the critical mass problem**: The benefits of a UPC system would not accrue to retailers, grocery manufacturers and hardware/software vendors equally. Retailers were the immediate benefactors in terms of cost savings while the benefits to manufacturers would come later – after the mass adoption by retailers. The benefits to retailers would be significantly reduced if the manufacturers' adoption remained muted. Hardware/software vendors also needed assurance that a minimum level of sale will be realized before investing in a new technology. The Ad Hoc Committee played a substantial role in the process by providing direction, opening channels of communication with the industry and using publications to sustain the momentum gathered. The committee was comprised of influential, well-respected executives from all areas of the industry supply chain so that all parties were well represented. The members of the committee leveraged their knowledge and experience in promoting the UPC system. ▶ **Separating the code and symbol**: Any proposal of the Ad Hoc Committee on the symbol of the UPC would have invited intense scrutiny from the industry. Thus, the committee decided to arrive at the code first and postpone the decision about the symbol. Any variation in length, structure and compatibility of the code dramatically altered the potential benefits to various participants in the supply chain. As a result, the committee realized that the decision on the code had to be sub-optimal in order to facilitate adoption from all parties. That would also ensure that the implementation costs were equally divided among the retailers and manufacturers. The committee also kept the interest in UPC alive through its publication, Progressive Grocer. The publication kept the industry abreast with the latest developments, timelines, addressing the interests of other parties such as consumers and trade unions and highlighting the advantages of the UPC system.

Coalition for Innovative Media Measurement contact:
Jane Clarke, Managing Director, CIMM
+1 212 590 2430
jane.clarke@cimm-us.org

Ernst & Young Global Media & Entertainment Center contacts

	Telephone	Email
Global Media & Entertainment Center		
John Nendick, Global Sector Leader (Los Angeles, US)	+ 1 213 977 3188	john.nendick@ey.com
Sylvia Ahi Vosloo, Marketing Lead (Los Angeles, US)	+ 1 213 977 4371	sylvia.ahivosloo@ey.com
Karen Angel, Global Implementation Director (Los Angeles, US)	+ 1 213 977 5809	karen.angel@ey.com
Matt Askins, National Accounting M&E Resident (New York, US)	+ 1 212 773 0681	matt.askins@ey.com
Yooli Ryoo, Knowledge Lead, (Los Angeles, US)	+ 1 213 977 4218	yooli.ryoo@ey.com
Jennifer Weiker, M&E Sector Resident (Los Angeles, US)	+ 1 213 977 3916	jennifer.weiker@ey.com
Global Area Leaders and Advisory Panel Members		
Farokh T. Balsara (Mumbai, India)	+ 91 22 4035 6550	farokh.balsara@in.ey.com
Mark Besca (New York, US)	+ 1 212 773 3423	mark.besca@ey.com
Neal Clarance (Vancouver, Canada)	+ 1 604 648 3601	neal.g.clarance@ca.ey.com
Yuichiro Munakata (Tokyo, Japan)	+ 81 3 3503 1100	munakata-ychr@shinnihon.or.jp
David McGregor (Melbourne, Australia)	+ 61 3 9288 8491	david.mcgregor@au.ey.com
Gerhard Mueller (Munich, Germany)	+ 49 891 4331 13108	gerhard.mueller@de.ey.com
Bruno Perrin (Paris, France)	+ 33 1 46 93 6543	bruno.perrin@fr.ey.com
Michael Rudberg (London, England)	+ 44 207 951 2370	mrudberg@uk.ey.com
Global Service Line Leaders and Advisory Panel Members		
Mark J. Borao, Global M&E Advisory Services Leader	+ 1 213 977 3633	mark.borao@ey.com
Thomas J. Connolly, Global M&E Transaction Advisory Services Leader	+ 1 212 773 7146	tom.connolly@ey.com
Alan Luchs, Global M&E Tax Leader	+ 1 212 773 4380	alan.luchs @ey.com
Advisory Panel Members		
Howard Bass	+ 1 212 773 4841	howard.bass@ey.com
Glenn Burr	+ 1 213 977 3378	glenn.burr@ey.com
Jonathan Dharmapalan, Global Telecommunications Leader	+1 415 894 8787	jonathan.dharmapalan@ey.com
Ian Eddleston	+1 213 977 3304	ian.eddleston@ey.com
Pat Hyek, Global Technology Leader	+ 1 408 947 5608	pat.hyek@ey.com
Bud McDonald	+ 1 203 674 3510	bud.mcdonald@ey.com
Andrew Sale	+ 1 213 977 3061	andrew.sale@ey.com
Ken Walker	+ 1 805 778 7018	kenneth.walker@ey.com

Ernst & Young

Assurance | Tax | Transactions | Advisory

About Ernst & Young
Ernst & Young is a global leader in assurance, tax, transaction and advisory services. Worldwide, our 141,000 people are united by our shared values and an unwavering commitment to quality. We make a difference by helping our people, our clients and our wider communities achieve their potential.

Ernst & Young refers to the global organization of member firms of Ernst & Young Global Limited, each of which is a separate legal entity. Ernst & Young Global Limited, a UK company limited by guarantee, does not provide services to clients. For more information about our organization, please visit www.ey.com

About Ernst & Young's Global Media & Entertainment Center
Whether it's the traditional press and broadcast media, or the multitude of new media, audiences now have more choice than ever before. For media and entertainment companies, integration and adaptability are becoming critical success factors. Ernst & Young's Global Media & Entertainment Center brings together a worldwide team of professionals to help you achieve your potential – a team with deep technical experience in providing assurance, tax, transaction and advisory services. The Center works to anticipate market trends, identify the implications and develop points of view on relevant industry issues. Ultimately it enables us to help you meet your goals and compete more effectively. It's how Ernst & Young makes a difference.

© 2011 EYGM Limited.
All Rights Reserved.

EYG No. EA0050
1104-1251132 L.A.

This publication contains information in summary form and is therefore intended for general guidance only. It is not intended to be a substitute for detailed research or the exercise of professional judgment. Neither EYGM Limited nor any other member of the global Ernst & Young organization can accept any responsibility for loss occasioned to any person acting or refraining from action as a result of any material in this publication. On any specific matter, reference should be made to the appropriate advisor.