EIDR for Sports Broadcasting Content Management Challenges

Richard W. Kroon
Entertainment Identifier Registry Association (EIDR)

Raymond Drewry
Motion Picture Laboratories (MovieLabs)

Abstract - *Any broadcast organization that remains static will soon become irrelevant. To remain competitive, broadcasters must constantly work to increase process velocity, accuracy, and flexibility while reducing time to market, manual touch-points, and associated labor costs. A necessary element in this is globally unique and persistent works identification, such as EIDR (Entertainment Identifier Registry) content and video service identifiers. Sports programming, as a broadcast case study, encompasses all of the issues found in the traditional broadcast supply chain, plus those common to other types of live events and additional complications unique to sports programming. To receive maximum benefit from standardization, a broadcaster must review its propriety processes, often built up over decades, to determine where and how to accommodate change. Effective solutions developed in the crucible of sports broadcasting have important workflow implications for all broadcasters, with particular relevance for those who work with non-sports live events and local origination programming.*

CONTENT MANAGEMENT CHALLENGES IN SPORTS BROADCASTING

Broadcasting is an unforgiving, real-time environment where the show must always go on – dead air is simply not an acceptable option. Live event broadcasting, including news, raises the ante by eliminating the delay between content creation and public presentation. Sports broadcasting is more challenging still, combining all the normal broadcasting issues (asset management for both acquired and locally-produced content, collating and presenting guide data, audience measurement, language versioning, VOD and ancillary digital distribution channels beyond initial broadcast, content archiving and retrieval, contract compliance, etc.) with the immediacy of live event programming (variability in start times and durations, weather and other delays, groups as participants, late changes to participants, etc.) and the special nature of sports programming (local blackouts, regional presentations, multiple productions of the same competition, frequent repetition of past matchups, frequent reuse of clips, sharing clips and content among broadcast competitors, etc.).

To remain financially competitive, all broadcasters must undertake continuous process improvement by increasing process velocity, accuracy, and flexibility while reducing time to market, manual touchpoints, and associated labor costs. One of the most effective ways to do this is by automating repeatable processes. Often, a necessary first step is standardizing manual processes so that they become amenable to automation. Along the way, organizations must also give up a certain amount of proprietary customization in the name of standardization, automation, and overall process efficiency. The limitations of these proprietary constructs, developed within broadcast organizations over time, are often not apparent until one attempts to improve workflows that include an external input or output. For workflows related to content management, these external interfaces almost always include manual touchpoints for title matching, which must be removed if the benefits of automation are to be realized.

GLOBALLY UNIQUE IDENTIFICATION

Traditionally, when two parties in the broadcast ecosystem have communicated about a work of common interest, they have identified the work by exchanging descriptive metadata along with any contract, query, asset, or transaction and then relied on manual labor to match everything together. Unfortunately, communicating parties do not always agree on what these descriptive data are or how they are structured – even titles may differ, especially when abbreviated to fit within arbitrary space limitations or when dealing with international releases or foreign works. Thus, record matching has been a long-standing challenge and remains an ongoing expense. This process can be improved significantly if the parties agree on a single identifier for each abstract work or individual asset [1].

Any asset identification scheme that is clearly defined and consistently applied can be used within a broadcast organization. In fact, organizations tend to have several different identification schemes in play at any given time, including separate IDs for accounting, production, scheduling, etc. When communicating with third parties, internal identifiers are of limited use. Organizations can agree on an identification scheme with each of their supply chain partners, but this leads to an explosion of point-to-point identifiers. In theory, there could be 10 different identifiers per title in a 5-party ecosystem, though the actual number is lower thanks to the use of commonly exchanged IDs, usually those of a dominant partner. Regardless of the number, the receiving parties must still manually match each

received ID. This could lead to five matching efforts per title in this example (the issuing party does not have to match the ID, having created it). Using a common, global identifier for these third-party interactions reduces the number of IDs in play to one, regardless of how many partners are involved in the exchange [2]. Of course, complex workflows do not have just one title exchange each, so the actual number of touchpoints and matching efforts per workflow can be quite high, with costs to match.

Cognizant Technology Solutions reviewed one such use case, where Warner Bros. and Microsoft adopted global identification for the delivery of feature films for Xbox Live playout. They calculated an immediate savings of 650 labor hours/year, with the potential to save 8,750 labor hours per year with 5 participating partners [3]. Broadcast workflows involve orders of magnitude more titles than the feature film use case studied, and broadcast supply chains involve far more than just two partners in their end-to-end workflows, so the potential for labor savings is significantly higher than calculated here.

There are several possible globally unique identification schemes that could be employed in the broadcast supply chain, including:

1. Classification systems, such as used by most libraries
2. Proprietary inventory indexing, such used internally by most broadcasters
3. Statistically-unique identifiers, such as UUIDs [4] and UMIDs [5]
4. Globally-unique, curated identifiers, such as the Content IDs provided by the Entertainment Identifier Registry Association (EIDR) [6]

Each type of identifier has its advantages, but also its limitations. Classification systems rely on human judgment to assign IDs, so there is no guarantee that the same work will get the same ID every time. The IDs can be parsed to extract the work's classification, but unambiguous identification still requires a shared resolution system. Indexing systems are easy to construct, but are only valid within their local domains, limiting both their coverage and their utility for multi-party communications without the burdensome cost of multiple manual matching touch-points. Statistically unique identifiers are also easy to produce, but they cannot be used by themselves to link related assets and do not come with a discovery mechanism that can resolve an ID to its descriptive metadata. In addition, it is not possible to say for certain that they will always be unique, just that it is very unlikely that they will duplicate [7]. For multi-party workflows and applications, the best overall solution is the use of a curated identifier that is guaranteed to be unique, persistent, and publicly resolvable, such as the EIDR Content and Video Service IDs [1].

EIDR AS A CURATED IDENTIFIER FOR BROADCAST

EIDR provides curated global registries of unique and persistent identifiers for audiovisual content (Content IDs) and delivery channels (Video Service IDs) based on the ISO/IES 26324 Digital Object Identifier (DOI) standard [8]. The International DOI Foundation guarantees the persistence and resolvability of all DOIs, so once issued, EIDR IDs are never deleted and can always be resolved to the associated record.

EIDR IDs are opaque, randomly assigned "dumb" numbers. Content IDs come from a pool of 1.2 septillion values (16^{20} = 1,208,925,819,614,630,000,000,000), while Video Service IDs come from a pool of 4.3 billion values (16^{8} = 4,294,967,296). EIDR IDs cannot be parsed to provide descriptive information. Instead, they are publicly resolvable, so with an EIDR ID, one can look up its descriptive metadata without restriction or charge. The reverse is also possible, using descriptive metadata to look up an associated EIDR ID [6].

EIDR Content IDs also can be retrieved using one of the many alternate IDs recorded in the EIDR Registry, such as an ID assigned to an asset by a production company, data aggregator, or audience measurement service. Additionally, the EIDR ID can be used as a pivot point, starting with one third-party ID and using that to retrieve a different third party ID from the EIDR Registry [9].
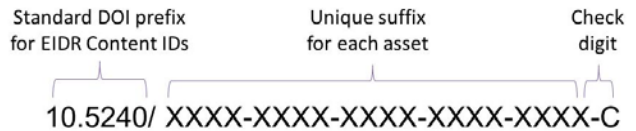
I. EIDR Content IDs

FIG 1 THE STRUCTURE OF AN EIDR CONTENT ID.

EIDR Content IDs support a hierarchical association, including works in the abstract (title records), versions of works (edits), and viewable representations of works (encodings or manifestations). The EIDR Registry contains metadata and APIs that allow traversal of the hierarchy and discovery of related items through relationships. This model works well with both traditional applications and the evolving world of linked open data (LOD) and the Semantic Web [10].

In the diagram that follows, the object names inside the boxes are those most familiar to broadcasters, while those above the boxes are the EIDR equivalents.
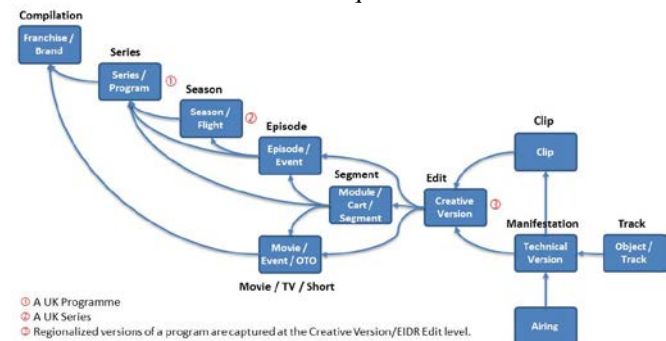

FIG 2 THE EIDR CONTENT ID HIERARCHY

In addition to the standard broadcast hierarchy, which accommodates both stand-alone works (made-for-TV movies, specials, etc.) and episodic works, the EIDR Content ID Registry supports multiple types of asset groupings and relationships, such as franchises, compilations, and composites. There are also a number of non-hierarchical relationships that allow one to specify other types of asset relationships, such as linking a promotional piece to the program it promotes, identifying one asset that represents alternate content for another, etc.

EIDR does not include presentation date and time information beyond the original showing of each work, so it cannot be used by itself as an airing ID. However, the Content ID (identifying an audiovisual asset, optionally at the title, edit, or encoding level) plus the Video Service ID (identifying a video delivery channel) can be combined with date/time information to identify a specific airing.

Once assigned, EIDR Content IDs can be carried with media files in metadata sidecars or other loosely coupled associations, or embedded directly in media files and data streams in a variety of ways, including [11]:

- **Standard Text Representations** – SMPTE Recommended Practice RP 2017 provides for standardized representations of EIDR IDs in text, KLV (key-length-value), and URI representations [12]. EIDR additionally defines a variety of representations including full and compact binary, Base64URL, and URN [13]. These representations can be used in/carried by multiple packages and standards, including XML.
- **MXF** – SMPTE is presently developing an EIDR Descriptive Metadata Scheme (DMS) to allow EIDR IDs to be carried in an MXF (Material Exchange Format) file [14].
- **AS-03 and AS-11** – The AMWA AS-03 (Program Delivery to multiple broadcast stations) and AS-11 (Program Contribution to a playout facility) subsets of MXF provide direct mechanisms for embedding EIDR IDs via the `AS_03_Identifier` and `Other Identifier` elements, respectively [15].
- **IMF** – EIDR IDs can be included in an IMF (Interoperable Master Format) Composition via the `//CompositionPlaylist/ContentVersion/Id` element [16].
- **BXF** – EIDR IDs can be carried in a BXF (Broadcast Exchange Format) data exchange wherever BxfAlternateId is used [17].
- **Open ID Binding to Essence** – A SMPTE study group is exploring ways to bind an EIDR ID to a media essence via audio and video watermarks [18].
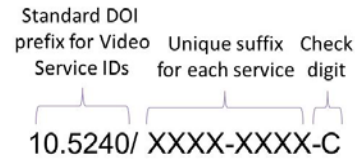
II. EIDR Video Service IDs

FIG 3 THE STRUCTURE OF AN EIDR VIDEO SERVICE ID.

EIDR Video Service IDs are a generic video distribution network identifier, designed to identify linear and non-linear audiovisual content delivery channels. They were developed to satisfy the requirements of the CableLabs Universal Content Provider Identifier [19].

The primary use cases for Video Service IDs cover the unique identification of traditional terrestrial, satellite, and cable linear broadcast channels and networks, such as ABC, HBO, ESPN, etc.; their related SD, HD and other variants; and their related regional or local feeds, such as East vs. West Coast and other regions that may receive different feeds for different reasons. In addition, Video Service IDs can identify online, IPTV, and VOD services.

These various delivery channels do not necessarily correspond directly to corporate entities. For example, Fox Sports West is not a separate company in its own right, but their corporate affiliations can be recorded. For example, ESPN, the channel, may be affiliated with ESPN (the company) and/or with The Walt Disney Company.
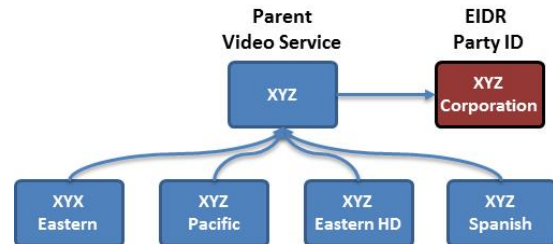

FIG 4 EIDR VIDEO SERVICES

In the preceding Video Service example, The XYZ channel has its own EIDR Video Service IDs as do its localized feeds, each of which carries the same programming but with specific variations in time offset, format, or primary language. The Party ID that defines the corporate parent for the XYZ channel is from another of the EIDR Registries, this time providing unique company identifiers, each of which begins with a 10.5237 prefix (to distinguish it from the 10.5239 prefix used by Video Service IDs) [9]. The following example shows how a Video Service can have more than one associated corporation (EIDR Party ID), which can be used to group Video Services by associated brand, while the example after that shows how to identify station affiliation using a combination of EIDR Party and Video Service IDs.
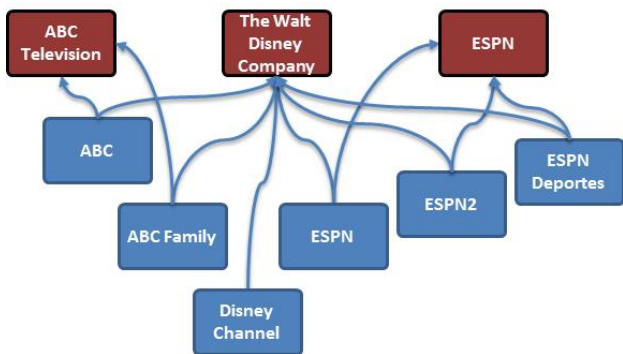
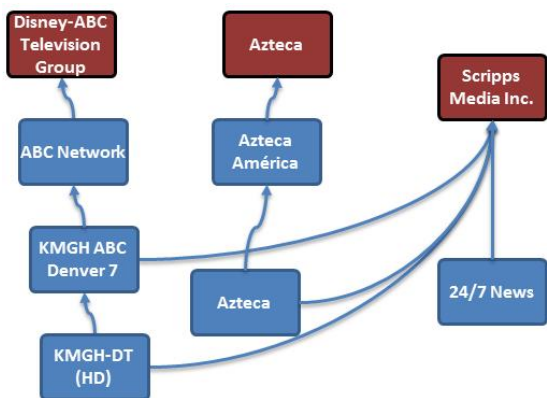FIG 5 USING EIDR PARTY IDS TO GROUP VIDEO SERVICES BY BRAND

FIG 6 USING EIDR IDS TO SHOW STATION AFFILIATION

There are two principal ways to identify regionalized content with Video Service IDs:

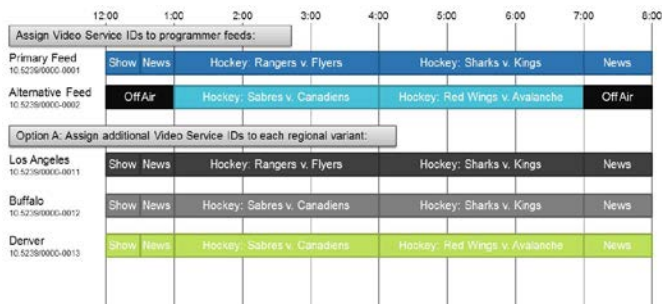1. Using Video Service IDs to identify each regional channel variant

FIG 7 EIDR VIDEO SERVICES WITH REGIONALIZED CONTENT, OPTION 1

2. Using Video Service IDs to identify the distinct video feeds carried with each channel
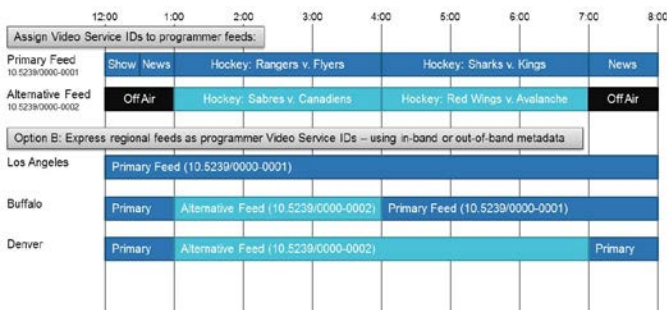

FIG 8 EIDR VIDEO SERVICES WITH REGIONALIZED CONTENT, OPTION 2

In the first option, regional variants of video services are created to represent each unique combination of actual content that may be delivered to a client, following the model typical for EPG data. Since, like Content IDs, EIDR Video Service IDs are durable and permanently resolvable, they are not used to identify ephemeral entities. So, just as a broadcaster does not change their call sign or channel name during the interval of a one-off event blackout, under this first option, one would not change the distribution channel's Video Service ID.

In the second option, the actual feed carried by each channel is identified by its own Video Service ID. There are no specific identifiers for blackout windows, but the blackouts can be identified in a particular channel by the switch to an alternate feed. While convenient for the sports broadcaster, this breaks the traditional EPG model, limiting its utility in many of the workflows with external touchpoints described in the section that follows [20].

The first option, where the content switch in a blackout is identified based on the Content ID of the presented program, rather than by the Video Service ID of the carried feed, is therefore preferred.

III. Broadcast Workflows with External Touchpoints

When combined, Content and Video Service identification can be used in most broadcast workflows with external touchpoints – the ideal use cases for curated identifiers. For acquired content, Content IDs arrive with the avails from the content producer. For locally produced content, the Content ID can be used as a project ID to coordinate amongst the various production partners and contractors. From there, EIDR Content IDs carry through the order and fulfilment process, possibly in combination with Video Service IDs for organizations that operate multiple channels. The IDs then flow through the broadcast organization and come out the far end in schedule data going to guide providers and audience measurement services. In the financial reconciliation sphere, Content and Video service IDs are used when communicating with music rights and retransmission collectives as well as the original content providers.
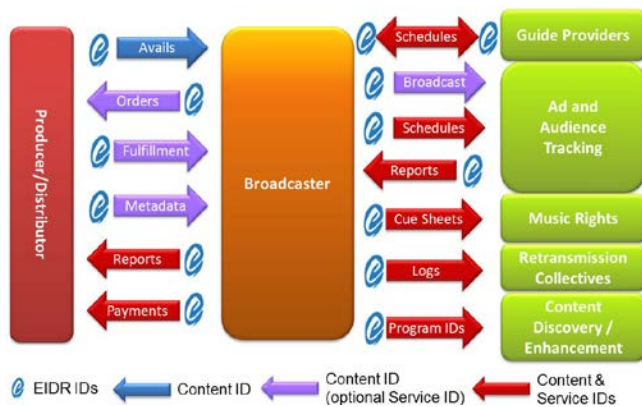
FIG 9 EIDR IDs IN WORKFLOWS WITH EXTERNAL TOUCHPOINTS

When communicating with partners who have not yet adopted EIDR, or when interfacing with legacy systems that cannot carry EIDR IDs, the EIDR alternate ID repository can bridge the gap. Each EIDR Content record can hold an unlimited number of alternate IDs, allowing immediate integration with any existing partner, workflow, or system.

THE SPECIAL CHALLENGES OF SPORTS BROADCASTING

I. Identifying the Identified

The first point to consider when using a curated identifier in sports broadcasting is what the identifier identifies. The possibilities include:

- A sporting event
- An audiovisual production of that event
- A content edit of that production of that event
- An encoding or other technical version of that edit of that production of that event

The correct choice depends on the workflow, since different situations require different levels of identification, singly and in combination.

Once what the identifier identifies has been determined for a particular workflow, the next point of consideration is the relationship between the identifier and the thing being identified. In most workflows, they are one and the same. In others, the relationship is defined by context and accompanying metadata. For example, in an avails situation, the acquired item is identified in an order by an Edit ID. The deliverable will be a technical representation of that Edit (an EIDR Manifestation), but it may still be associated with the EIDR ID of the order. The context of the order, metadata accompanying the delivery, or simple inspection upon receipt can identify the nature of the technical version actually delivered (ProRes vs. MPEG, SD vs HD, etc.), so an EIDR Manifestation ID is not necessary.

Within the broadcaster, the "original," or Grand Master, version of a program has its own EIDR Edit ID, but derived versions (syndicated cut downs, time tailored versions, international versions, etc.) need not if they are sufficiently

well-defined by context and accompanying metadata. If there is any chance that they might escape into the wild without their qualifying characteristics, then certainly a separate Edit ID should be assigned. The same is true for technical versions and encodings. The master mezzanine should have its own Manifestation ID, but derived encodings may not if they can be clearly, consistently, and unambiguously identified from context or inspection. As with Edit IDs, when in doubt, identify by assigning a unique Manifestation ID to the derived versions.

These issues are common to all broadcasters. Sports broadcasting simply raises the ante by adding the notion of the event that is being presented and placing it on top of the normal hierarchy. EIDR IDs cannot be used to identify events, just as they cannot be used to identify the source material for a motion picture or television episode, but they can carry alternate identifiers that do. So, if an organization has an event identifier, it can be added to the EIDR registry so the programs can be tied to the events they depict. Since different organizations may use different event identifiers, the alternate IDs in EIDR carry a domain value to identify the authoritative source of the identifier. For example, Flixster/Rotten Tomatoes Content IDs have a "flixster.com" domain to distinguish them from other IDs, such as ITV IDs, which have an "itv.com" domain.

Following is a list of the recommended EIDR ID associations:

- Production project identifier → Title ID
- Avails and Acquisitions → Edit ID
- Content Delivery → Varies by delivery mechanism, but generally matched to the order ID with the option to include a Manifestation ID
- Playout and Traffic Management → Manifestation ID
- Asset and Archive Management → Manifestation ID (perhaps grouped by Title or Edit ID)
- EPG and Content Discovery → Edit ID
- Ad and Audience Tracking → Title ID
- Music Rights and Broadcast Logs → Edit ID

II. Clips and Granular Identification

Sports broadcasters are uniquely clip-happy. Programs are regularly sliced up into a variety of clips that are then re-purposed in innumerable ways. They are also shared with other broadcasters (who did not carry a program depicting a particular event) in a reciprocal exchange for other clips. They appear in promotional bumpers and interstitials and are used in news, commentary, and highlight shows that recap games, summarize activities (such as the "Hit of the Week"), and recall the past times when something similar took place.

An editor may cobble together clips from a half-dozen different sources, mixing in-house and acquired content, and present them as a single package without recording the original source of the clips. Since digital media does not suffer from analog generational loss, editors no longer have to go back to the source material for their clips, and may clip from previous packages of clips or other shows that used

clips from other sources. Unraveling all of this after the fact can be nearly impossible if not cost prohibitive.

These clips are everywhere, and they are very poorly identified and tracked. This can quickly become a contract compliance issue if there are any limitations on the use of a particular clip, if per-use royalties are owed, or if source acknowledgements are required.

To address these issues, EIDR provides a clip identifier. It is a Content ID that has an Edit or another Clip record as its parent and can identify the time-offset duration of the material extracted into the new clip. Because these are standard EIDR Content IDs, they can be embedded directly in media files, ensuring that the Clip ID travels along with the clip wherever it goes. (From the Clip ID, one can resolve the Edit ID and ultimately the Title ID of the show from which the clip was extracted.)

The question then becomes, just because you can identify a clip does that mean that you must identify a clip? This harkens back to the section on "Identifying the Identified" and the discussion of the association between an ID and the thing being identified. In general, one should clearly identify down to the level of the "just reported difference," a concept similar to the "just noticeable difference." In this case, one provides a unique identifier for anything that needs to be individually reported. That is why traffic management is at the Manifestation level while music rights payments are at the Edit level – it matters if the content you play on air is the SD or HD encoding of a program, but the music rights do not vary at the Manifestation level, so reporting Manifestation IDs with a music cue sheet is unnecessary. Cue sheets can be tied to Title IDs, but that runs the risk of reporting incorrect results when edits use different music.

So, in a particular broadcast application, does a clip represent a just reportable difference? Are there royalties that must be paid based on how often and where a clip is presented? Is the Research Department collecting audience performance data at the clip level? The likely answer today is, "No." If the reasonable answer for tomorrow is, "Yes," then there may be motivation to start recording at the clip level now to be future proofed for when the need arises.

The exception to this are clips that carry license restrictions or that are prepared for individual Web streaming. In the first case, if usage royalties must be paid (or will be earned) on a licensed clip, then it is important that its uses be closely tracked. The sports broadcaster will need to have a tracking mechanism regardless. Adding an EIDR Clip ID to the tracked asset is a simple matter. In the second case, if a contiguous program clip is presented as a single unit that a consumer may view, then it makes sense to register it so it can be tracked and reported going forward. (A package of clips assembled from different programs or non-contiguous segments of a single program is registered as a unique composite program with a unique Title ID or Edit ID, respectively.)

III. Variable and Inconsistent Descriptive Metadata

To assign an ID, one must be able to describe the work being identified with sufficient clarity to distinguish it from all other similar works. The same is true for edits and manifestations, but there the de-duplication domain is constrained to sibling edits of the same abstract work or sibling manifestations of the same edit, so clear and differentiating metadata is much easier to come by. This is an issue for all ID registrants, but it is of particular concern in sports broadcasting.

The primary distinguishing characteristic for most broadcast programs is the title (for episodic programs, the episode title within the scope of the series or season). In sports, event titles are often based on the names of the competitors in the sporting event. This becomes problematic when:

- One event has multiple sub-events (as with a track and field competition, gymnastics, ice skating, etc.)
- One event has multiple competitors (as with tournaments, the undercards on a boxing event, the Olympics, etc.).
- A sporting event spans multiple days (as with the Tour de France, the World Cup, Wimbledon, etc.)
- The same competitors meet multiple times in the same day (as with a baseball double-header – there hasn't been a triple-header since 1920 [21])
- An event has last-minute competitor changes (due to illness or injury and as may be the case in a single-elimination tournament)
- An event is re-scheduled from its original date/time (due to weather or other external factors).
- The same scheduled program carries different events in different regions (as when the College Football Game of the Week features different games depending on the local conference), blurring the line between series and brand/franchise.
- Multiple programs are produced depicting the same event – these may be produced by the same company for different markets or produced by different companies. For example, for the recent London Olympics, the BBC broke out individual events like the 100M final into their own programs, while ITV presented large 3-hour blocks of Olympics coverage featuring the same events.
- Multiple programs are produced using the same raw materials. Using the Olympics as an example, which is perhaps the most complicated sporting event to host or to broadcast, the host country provides the same feeds to all broadcasters, who then produce their own programs. Formula One racing also provides a common feed, but individual broadcasters may augment this with additional material focused on a particular car or driver.
- The same program is presented with different primary languages. This is akin to dubbing a scripted program and may come with or without language-specific lower-thirds and motion graphics

- The same program is simulcast on multiple distribution channels or media (for example, a TV/radio simulcast, a broadcast/Web streaming simulcast, or broadcast by multiple outlets with local station identification).

In addition, most broadcaster master title management systems track the actors who appear in scripted fiction programs, helping to identify and disambiguate different programs. (There is some debate if "reality TV" programs are non-fiction, but regardless, their participants are treated as actors in most title management systems.) For sporting events, the competitors are rarely, if ever, captured directly and may only appear in the title or synopsis. Directors are also of significance in title de-duplication. They are generally available for motion pictures, often available for scripted episodic, but hardly ever available for live events, news, and sports.

So, how do sports broadcasters identify their programs so they can be differentiated from one another and remain discoverable should a metadata search be required to extract an assigned ID? The three most useful identifying characteristics for a sports program tend to be:

- **The Title** – Almost any naming convention can be used so long as it is consistently applied and is sufficient for human disambiguation. Organization of registrations into episodic structures, where applicable, can help disambiguate similar titles.
- **The Event Schedule** – While EIDR does not provide an Event ID, it can track the date and time (with local time zone) of the scheduled start of an event. This can help differentiate similar programs presented on the same day. (Of course, if one has a unique event ID, it can be added to the EIDR record as an alternate identifier.)
- **The Program Producer** – When the same event is covered by multiple broadcasters, title and schedule are insufficient to differentiate their individual productions. The production company (or companies), and optionally the primary broadcaster (a Party ID identifying the broadcast company, rather than a Video Service ID identifying the broadcast channel) for whom a program has been produced, can help disambiguate the programs.

Additional identifying characteristics are also useful to record, and should be included as part of a standardized registration practice, including:

- The original broadcast date (or anticipated broadcast date for a program that has not aired yet)
- Anticipated running time (useful to distinguish a full game from game highlights)
- The primary language used in the presentation (for sports, this is generally the language spoken by the commentators and use for on-screen text)
- Any alternate titles by which the program may be known (quite useful for language disambiguation and later discovery)
- Any alternate identifiers that may be available (including house production IDs and event IDs)

Certain accommodations can be made if not every data point is available in every situation. As a final stopgap, if the EIDR automated de-duplication system cannot say with confidence that a proposed program registration is unique (or matched to a previous record), the registration is referred for manual review. This adds a delay to the registration process, so it is always better to provide as much descriptive metadata as possible up front so the automated system can make its best determination. Of course, registering at the earliest practical point is also highly recommended. (Keep in mind that the de-duplication domain can be constrained by using an episodic structure and is naturally constrained for edits and manifestations, so not every registration runs the risk of requiring manual review.)

To address the variability inherent in sports broadcasting, and to account for the natural evolution of a program that is registered early, the initial ID registration must be followed by metadata updates to the EIDR Registry should any of the key identifying characteristics change prior to the program's initial presentation.

IV. Variable and Inconsistent Program Schedules

Traditional broadcast programming is only occasionally pre-empted by another program – and then generally only in the case of important breaking news. Pre-emption and cancellation are a regular occurrence in sports broadcasting, as they are in other live events broadcasting. Live events can also run long – and occasionally short – leading broadcasters to either switch to their next program early (leaving the audience for the next program to "join in progress") or to air alternative programming to fill the time slot. (Live post game analysis programs are particularly handy since their durations can be adjusted to re-align the broadcast schedule.)

There is another issue, unique to sports broadcasting, that confounds program schedules, and that is the mid-game switch, or what ESPN calls a "whip-around." In this case, a scheduled game is not as exciting as anticipated, so, to avoid losing current viewers, the broadcaster switches to a more interesting game.

EIDR Content IDs identify the program, not the events being depicted. If for some odd reason, *NFL Monday Night Football* presented celebrity curling with play-by-play by Chris Berman instead of football, the program would still be *Monday Night Football*. In the case of a program pre-emption or whip-around, there are two separate programs, where one replaces the other within the first program's scheduled time slot.

All of this causes consumer confusion, since program guide information published in advance will not match the programs actually being presented on air, and audience measurement issues, since ratings are generally reported according to the scheduled programs. (This is not an issue for music rights, retransmission royalties, and other financial reconciliations, since those take place far enough after the

airing that certified broadcast logs can clear up any confusion.)

The ultimate technological solution is to embed EIDR IDs in the broadcast stream, so guide providers, audience measurement firms, and others will know exactly which program is being aired on which channel at which time, and can adjust for any deviations from the published schedules [23]. Until that time, several standard practices limit the organizational impact of sports programming schedule vagaries.

- For guide and immediate audience measurement purposes, the EIDR ID identifies the program intended for presentation. This is tied to the time slot. (Recall from the earlier discussion that an EIDR ID identifies one thing but can be associated with another thing – here, it identifies the program and is associated with the time slot.)
- The broadcaster will still track the EIDR IDs of the programs actually presented (ideally, drawn from the traffic management and playout systems that delivered the content to air) and use that in its final reconciliations (including adjusting audience measurement data was reported by time slot, not by program).

Local pre-emptions and blackouts are a slightly different matter, since they are generally established in advance, though the lead-time may still be quite short, since it is generally dependent on the rate of ticket sales for the depicted event. They mostly apply to sports broadcasting, but may also affect concerts or other live events for which attendees pay an admission fee. The FCC's non-duplication rule can also come into play, where a local broadcaster has exclusive rights to a program also carried by a distant station being presented in the local area on a cable or satellite system [24]. In any case, these blackouts (and the programming content that replaces them) are addressed using a combination of Video Service and Content IDs, as noted earlier.

V. Deviations from Standard Episodic Structure

The challenges of traditional episodic program identification are well understood and easily addressed following EIDR's established best practices, including the usual edge cases of season-less series, strip shows, miniseries, re-cut series, and modified distribution bundles. Sports networks also have episodic programs, but because they are developed and presented in an environment optimized for sports programming, they do not have many of the back office structural elements normally associated with episodic programs. As a result, when sports broadcasters do produce episodic programs, they are not structured and managed in a way that would be recognized by traditional broadcasters or accommodated by most asset and metadata tracking systems. This provides a particular challenge when applying standard practices to sports network episodic programs. For example:

- When organizing sports into traditional broadcast seasons, how does one categorize playoffs, pre-season play, exhibition games, and other activities that take place outside the established season?
- In college sports, the conferences are often independent brands, so if one were to put college football into an episodic structure, would that require one series for all of college football or one series for each of the conferences? If you attempted the latter, where would you put an inter-conference game?
- Annual and quadrennial events, such as the Super Bowl and the Olympics could be individual series, but they are better served by collecting them as brands or franchises (under an EIDR Compilation) and dealing with their recurrences separately.
- Professional sports leagues and teams have their own broadcast outlets, often presenting the same material as the sports networks, but packaged in different ways.

These issues are not insurmountable, but they do require individualized solutions. While each broadcaster's mapping to the EIDR episodic structure may be slightly unique, it is still consistent within the organization. Sports broadcasters may need to customize their episodic mappings per sport, league, or show to accommodate these variations. Then, individual program registration can be automated so that an EIDR ID is created whenever a new record is defined in the broadcaster's internal title system.

This additional up-front investment is more than recouped through future savings, since each new entry in an established series or season enjoys a constrained de-duplication domain and is much less likely to bump up against a confusingly similar record that requires manual differentiation. Once sports broadcasters get down to edit, clip, and manifestation levels, there is no material difference between episodic and non-episodic programs and standard automation rules apply.

VI. Acquisition of Off-Air Recordings

This is not so much a special problem of sports broadcasting as an increasing opportunity for all broadcasters where curated identification can play an important role. Currently, off-air recordings may be captured for use as catch-up TV and VOD content. In addition, most television and radio archives record off-air programs and add them to their holdings as digital assets. This represents a high volume of content, both in terms of total hours of programming and in individual catalog items, so every economy of scale must be applied to make this acquisition activity practical.

These materials can be collected in digital form using automatable, scalable, and mature off-air capture platforms, such as Cambridge Imaging Systems' Orbital [25], but the descriptive metadata captured with the broadcast materials tends to be both low density and low quality. This then requires a separate acquisition step to obtain complete

descriptive metadata and a matching process to link this to the recordings.

Embedding EIDR Content and Video Service IDs for program and delivery channel identification into the broadcast stream [23], EPG data, certified broadcast logs, and descriptive metadata provided by the broadcasters themselves and third party data aggregators, allows for the collation of these disparate sources in an automated workflow. This produces a rich content record for down-stream activities, including VOD deliveries and archive client search and discovery. This process also adds value to the recordings themselves, making them more available to the original broadcaster or other licensees for reuse and resale across the broadcast sector [1].
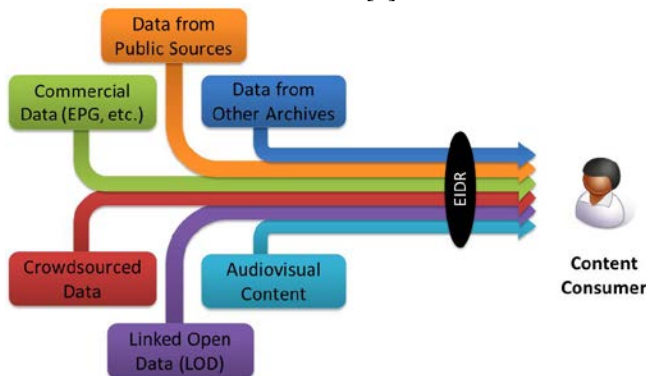

FIG 10 LINKING DISPARATE DATA SOURCES WITH AN EIDR ID

CONCLUSION

Sports broadcasters have issues. Their issues appear in greater number and variety than for other types of broadcasters, since they start with all of the complications of traditional broadcasting, including live events and news, and add to that the unique world of sports. This has led to the creation of sports-specific solutions, which occasionally get in the way as sports broadcasters evolve and standardize their operations to take advantage of new revenue and cost-savings opportunities. Many opportunities for improvement require the participation of supply chain partners outside the fixed walls of the sports broadcaster, leading to heavy reliance on manual title matching to facilitate these content management related workflows. Adopting global, shared identification using a curated registry such as provided by the Entertainment Identifier Registry (EIDR) for both content and delivery channel identification can help eliminate these manual touchpoints, reduce costs, and open up new opportunities for automation and process improvement. Other broadcasters can learn from these sports examples, as they adopt shared identification and work to align their internal systems and operations with their supply chain partners.

REFERENCES

[1] Kroon, R. W., et al. "Content Identification for Audiovisual Archives," *International Association of Sound and Audiovisual Archives Journal*, No. 45, Summer 2015.

[2] Kroon, R., "Bringing Order to Digital Identifiers," *M&E Journal,* Winter 2014, pp. 148-150.

[3] Drewry, R. and Dulchinos, D., "Transforming Entertainment through Technology," *M&E Journal,* Fall and Winter 2013, pp 81-88.

[4] ITU-T (Telecommunication Standardization Sector of ITU), "Information technology – Procedures for the operation of object identifier registration authorities: Generation of universally unique identifiers and their use in object identifiers," www.itu.int/rec/T-REC-X.667-201210-I/en. Last modified October 2012.

[5] Wallaskovits, N. and Liebl, C., "UMID – Unique Material Identifier." www.digitalpreservationeurope.eu/publications/briefs/UMID_Unique%20Material%20Identifier.pdf. Accessed November 16, 2014.

[6] Entertainment Identifier Registry Association, "About EIDR," eidr.org/about-us, Accessed January 24, 2015.

[7] Leach, P., M. Mealling, and R. Salz, "A Universally Unique IDentifier (UUID) URN Namespace," www.ietf.org/rfc/rfc4122.txt. Last modified July 2005.

[8] International DOI Foundation, "The DOI System," www.doi.org. Last modified January 21, 2015.

[9] Entertainment Identifier Registry Association, "EIDR System version 2.0 Registry User's Guide," eidr.org/documents. Last modified August 13, 2013.

[10] Entertainment Identifier Registry Association, "Overview," eidr.org/documents. Last modified July 2014.

[11] Lemieux, P-A, "The Use of EIDR in Broadcast Workflows," eidr.org/documents. Last modified January 2015.

[12] SMPTE, "Digital Object Identifier (DOI) Name and Entertainment ID Registry (EIDR) Identifier Representations," RP 2079:2013.

[13] Drewry, R., "EIDR: ID Format," http://eidr.org/documents/EIDR_ID_Format_v1.2.pdf. Last modified March 3, 2014.

[14] SMPTE, "Carriage of EIDR Identifiers in MXF files," TC-31FS DG RP 2089.

[15] Advanced Media Workflow Association, "Application Specifications," www.amwa.tv/projects/application_specifications.shtm. Accessed January 24, 2015.

[16] Chang, A., "SMPTE Standards PDA Webcast: IMF (Interoperable Master Format)," https://www.youtube.com/watch?v=bmhv36hmSP4. Last modified January 16, 2013.

[17] SMPTE, "Using Ad-ID and EIDR as Alternate Identifiers in SMPTE BXF and ATSC PMCP," RP 2021-5.

[18] SMPTE, "Open ID Binding to Essence," 24TB Study Group.

[19] Entertainment Identifier Registry Association, "Proposal for Video Service ID (Network ID)," eidr.org/documents. Last modified January 8, 2012.

[20] Agranoff, D., "Video Services," eidr.org/documents. Last modified April 2014.

[21] Suehsdorf, A. D., "The Last Tripleheader," Society for American Baseball Research - Research Journal Archives, research.sabr.org/journals/last-tripleheader. Accessed January 25, 2015.

[22] A&E Television Networks, "Nov 17, 1968: The Heidi Bowl," www.history.com/this-day-in-history/the-heidi-bowl. Accessed January 25, 2015.

[23] Coalition for Innovative Media Measurement, "TAXI – Trackable Asset Cross-Platform Identification," cimm-us.org/initiatives-2/taxi. Accessed January 24, 2015.

[24] Kroon, R. W., "program non-duplication," *A/V A to Z*, 2010, pp. 528-529.

[25] Cambridge Imaging Systems, "Orbital Enterprise Video Recording," www.cambridgeimaging.com/orbital_enterprise_video_recording. Accessed November 24, 2014.