CONTENT IDENTIFICATION FOR AUDIOVISUAL ARCHIVES

Richard W. Kroon, Entertainment Identifier Registry Association, USA
Raymond Drewry, Motion Picture Laboratories, USA
Andrea Leigh, Library of Congress, USA
Stephen McConnachie, British Film Institute, UK

1.   Abstract

Digital archives exist within a complex global web of interests and agencies with ever-increasing — and often conflicting — demands placed on their limited resources. Issues with work identification confront them during acquisition, reconciliation, and de-duplication of assets obtained from multiple sources. Furthermore, they must accommodate manual workflows and the resulting process delays, volume constraints, and costs of error correction. A curated system of unique, global identification for audiovisual works, their derivative versions, and physical and digital manifestations as provided by the Entertainment Identifier Registry Association (EIDR) can directly address many of these issues and enable the provision of systems and services that address many more. We identify several use case examples, including a federated system for archive search and retrieval, the collection of off-air television broadcasts, and addressing intellectual property rights assertions.

2.   Why Identify?

Archives do not exist in isolation; they are part of a wider ecosystem with each institution catering to a particular audience. The value of an audiovisual archive, however you choose to measure it, is always less than its value when part of a larger, interconnected network of archives.[32] Cooperation among archives helps fill the gaps in holdings and services that inevitably exist in any collection, while also providing opportunities for economies of scale, sharing costs, and collaboration on best practices. These advantages come at the cost of additional competition for an archive's already limited time and resources and sometimes may conflict with an individual archive's priorities, but the overall effect of collaboration is still significantly positive. A reliable method of shared identification minimizes costs and maximizes benefits: common assets are identified accurately; archives can easily exchange assets and metadata; federated search and cross-archive discovery become simpler; and digital delivery of assets becomes more streamlined. A system of globally unique, curated identifiers can also benefit an archive's internal operations, particularly during content acquisition, reconciliation of materials obtained from multiple sources, and asset digitization. Shared, curated identification also helps eliminate manual touch points in most workflows, leading to increased process accuracy and throughput while simultaneously reducing costs and errors. Many of these benefits mirror the improvements a shared, curated identifier has already brought to the commercial audiovisual sector.[33],[34]

Once you can clearly articulate what exists in a collection, the next hurdle is making that information available to those who need it: if the people who need you cannot find you, then you may as well not exist. The Internet has made searching easier — both searching for archives and searching an archive for items in its collection — but this is still far from the ideal. Not all archives are online and not all collections are equally accessible, leading to a continued reliance on word-of-mouth and oral history to identify the holders of desired assets.[35] At the same

---

32 The collective value of a network of archives follows the pattern established by Metcalfe's Law, with the network's value increasing at a greater rate than the simple sum of its individual members. ("Metcalfe's Law," http://www.businessdictionary.com/definition/Metcalfe-s-Law.html. Bob Briscoe, Andrew Odlyzko, and Benjamin Tilly, "Metcalfe's Law is Wrong," http://spectrum.ieee.org/computing/networks/metcalfes-law-is-wrong.).

33 Raymond Drewry and Don Dulchinos, "Transforming Entertainment Through Technology," *M&E Journal* (Winter 2013-14): 81-86.

34 "Streamlined Content Metadata Integration and Management Using Entertainment ID Registry (EIDR)," http://www.eidr.org/documents/Expo2012_Technical-PaperEIDR.pdf.

35 As evidenced by the regular research assistance requests on AMIA-L, the Association of Moving Image Archivists Listserv: http://www.amianet.org/participate/listserv.php.

time, digital media can allow simultaneous, geographically distributed access without degrading physical archival assets, but only for those archives that can afford to digitize their analog collections and also have the means to deliver them to their clients.

This is just a modern example of the problem of naming things. If there is no name for something, you cannot talk about it. If you are working with someone else, the two parties must have either a shared name or a way of translating each other's names. If neither of those is present, no exchange of information is possible. In that worst case, "Wovon man nicht sprechen kann, darüber muss man schweigen. (Whereof one cannot speak, thereof one must be silent.)"[36]

3.   How to Identify

A collection is only as good as its identification system: if you cannot find something, then you may as well not have it. Humans and machines search differently. People are very good at using imprecise search terms and finding the results that matter to them. Machines are more effective with precise searches. For example, principles for constructing preferred titles for motion pictures in libraries have largely been unapplied. In libraries, cataloging practice is predicated upon a reliance on transcription from a chief source of information, typically from information as it appears on the publication in hand. In moving image archives, description is reliant upon a preference for the work's origin and historical context. In other words, general library catalogs are built to support the discovery of specific publications and their various editions, whereas a catalog in a moving image archive is designed to differentiate among works and variants of works. This difference in interpretation is a significant reason why the principles for constructing preferred titles is inconsistent at best, as can be illustrated by the various versions of the motion picture Blade Runner (1982):[37]

- Blade runner (Computer file)
- Blade runner (Motion picture)
- Blade runner (Motion picture : Director's cut)
- Blade runner (Motion picture : Final cut)

The main purpose for the addition of the qualifier "(Motion picture)" is to both collocate and differentiate among the various motion picture versions of *Blade Runner*, as well as to differentiate the motion picture from other forms of the work, in this instance the 1997 video game qualified as a computer file. This reliance on human-readable text strings is highly subjective and cannot be easily understood without a human being referencing the entire catalog record.

The preferred title text strings also do not correlate to how moving image works are commonly referenced in other contexts, particularly foreign language versions registered by local archives in other countries, such as Venezuela (El cazador implacable) or Russia (Бегущий по лезвию).[38] Since libraries are not oriented towards researching the distribution history of a motion picture, the foreign language release title would be transcribed in a separate catalog record for the videorecording of the film, with the foreign language appended to the preferred title: Blade runner (Motion picture). Russian. This approach is useful for human-readable browse displays, but is not suitable for machine matching, as it requires that all databases conform to the same principles consistently.

---

36 Ludwig Wittgenstein, *Tractatus Logico-Philosophicus* (New York: Harcourt, Brace & Company, Inc., 1922), Proposition 7.

37   The preferred titles for *Blade Runner* can be located as part of the Library of Congress linked data service, available at http://id.loc.gov, or at Library of Congress Authorities, available at http://authorities.loc.gov. Instructions for the formulation of preferred titles are included as a Library of Congress-Program for Cooperative Cataloging Policy Statement as part of the toolkit for *Resource Description and Access* (RDA LC-PCC PS 6.27.1.9 Appendix 1), available at: http://access.rdatoolkit.org.

38 "Blade Runner (1982) Release Info," http://www.imdb.com/title/tt0083658/releaseinfo?ref_=tt_dt_dt#akas.

Traditionally, when two parties 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 of 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. However, this is a point-to-point solution and the number of identifiers expands geometrically with the number of parties involved. Worse yet, since archive management and other workflow support systems are not particularly flexible, there is often no place to record these point-to-point identifiers in each party's support systems, so they reside off to the side, requiring manual effort to re-link them at every touch point.

We cannot solve all of these issues today, but we can begin with clear and unambiguous identification of audiovisual works, encompassing moving images and radio programs. There are several possible identification schemes, including:

- Content classification, such as used by the Library of Congress[39]
- Proprietary inventory indexing, such as used at the British Film Institute[40]
- Statistically-unique identifiers, such as UUIDs[41] and UMIDs[42]
- Globally-unique, curated identifiers, such as the Content IDs provided by the Entertainment Identifier Registry Association (EIDR)[43]

Each type of identifier has its advantages, but also its limitations. Classification systems are widely used, but can be cumbersome to assign and are open to subjective interpretation. Indexing systems are only valid within their local domains, limiting both their coverage and their utility. Locally generated identifiers are easy to produce and statistically unique,[44] but they cannot be used to link related assets and do not come with a discovery mechanism that can resolve an ID to its descriptive metadata. For workflows and applications such as described above, the best overall solution is the use of a curated identifier, such as the EIDR Content ID.

EIDR IDs are globally unique; externally resolvable; applicable to works in the abstract (title records), versions of works (edits), and representations of works (encodings or manifestations); and are able to support multiple types of asset groupings and relationships and to store multiple alternate titles and alternate identifiers per asset.[45] This last feature starts to alleviate the point-to-point translation problem: EIDR acts as a bridge between multiple systems.

39 "Library of Congress Classification," http://www.loc.gov/catdir/cpso/lcc.html.
40 "Advanced search," http://collections-search.bfi.org.uk/web/search/advanced.
41 "Information technology — Procedures for the operation of object identifier registration authorities: Generation of universally unique identifiers and their use in object identifiers," http://www.itu.int/rec/T-REC-X.667-201210-I/en.
42 Nadja Wallaskovits and Christian Liebl, "UMID — Unique Material Identifier," http://www.digitalpreservationeurope.eu/publications/briefs/UMID_Unique%20Material%20Identifier.pdf
43 "About EIDR," http://eidr.org/about-us.
44 It is not possible to say for certain that they will always be unique, but it is very unlikely that they will duplicate. (P. Leach, M. Mealling, and R. Salz, "A Universally Unique IDentifier (UUID) URN Namespace," http://www.ietf.org/rfc/rfc4122.txt.)
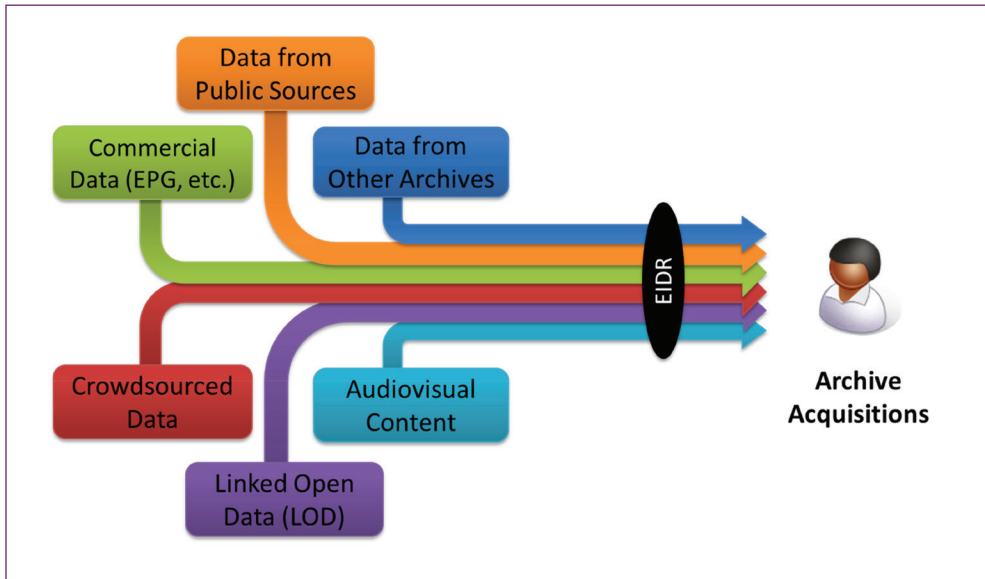45 "Overview," http://eidr.org/documents/2014-07-01-EIDR_Overview.pdf.

Figure 4. Bringing disparate elements together for an archive using a shared EIDR ID.

The ID itself is an ISO-standard DOI (Digital Object Identifier),[46] which, despite its name, can be applied to physical or digital assets.[47]
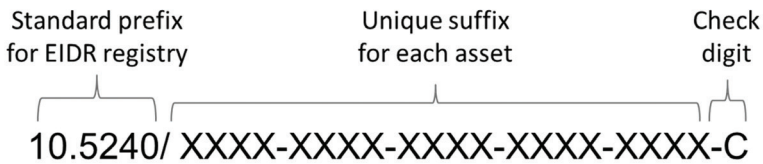

Figure 1. The structure of an EIDR Content ID.

EIDR IDs are opaque, "dumb" numbers randomly assigned from a pool of 1.2 septillion possible values.[48] They cannot be parsed to provide descriptive information as with a classification scheme. Instead, they are publicly resolvable, so if you have an EIDR ID, you can look up its descriptive metadata without restriction or charge. You can also do the reverse, where you use the descriptive metadata to look up the ID. Alternatively, if you have one of the many alternate IDs recorded in EIDR records, such as an ID provided by a data aggregator, another archive, or a production company, you can use that ID to retrieve the associated EIDR ID or pivot on the EIDR ID and retrieve one of the other alternate identifiers, depending on your particular need.[49] EIDR IDs are also durable: once issued, they are never deleted, so they can always be resolved to the associated content record.[50]

---

46 "Key Facts on Digital Object Identifier System ," http://www.doi.org/factsheets/DOIKeyFacts.html.
47 A Digital Object Identifier is a digital identifier of objects, not an identifier of digital objects. This is akin to the confusion surrounding the one-eyed, one-horned, flying, purple people eater, which is, in fact, an eater of purple people not, as so many incorrectly assume, a purple eater of people. (Sheb Wooley, "The Purple People Eater" (New York: MGM, 1958).)
48 $16^{20} = 1,208,925,819,614,630,000,000,000$.
49 "EIDR System version 2.0 Registry User's Guide," http://eidr.org/documents/EIDR_2.0_Registry_User_Guide.pdf.
50 If an EIDR Content record is deleted, the EIDR ID is aliased so that it resolves to the EIDR Tombstone record. The International DOI Foundation guarantees the persistence and resolvability of all issued DOIs.

The hierarchical EIDR record examples are based on a theatrical motion picture and a television series, but the same hierarchical registration structure can be applied to audio-only works, such as radio programs,[51] making EIDR Content IDs equally applicable to the management of both moving images and radio programs.
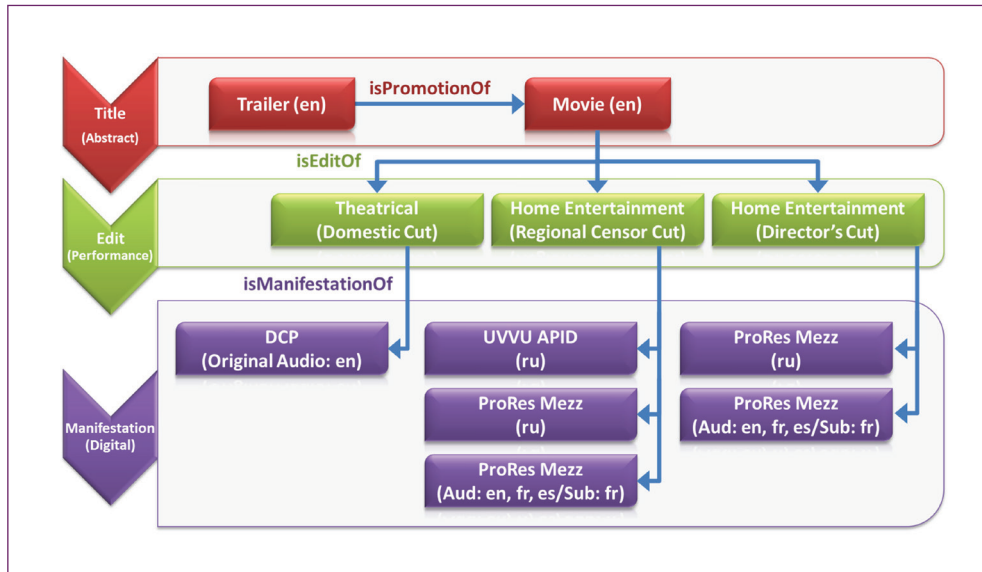

Figure 2: Sample EIDR registration tree for stand-alone works.[52]

51 Music recordings are best identified with an ISRC (International Standard Recording Code, www.usisrc.org). Music videos can be identified with either or both an ISRC or EIDR ID.
52 Commercials and other advertising ephemera are best identified with an Ad-ID (www.ad-id.org). Trailers (as shown in the diagram) and other items that promote audiovisual works can be identified with either or both an Ad-ID or EIDR ID.
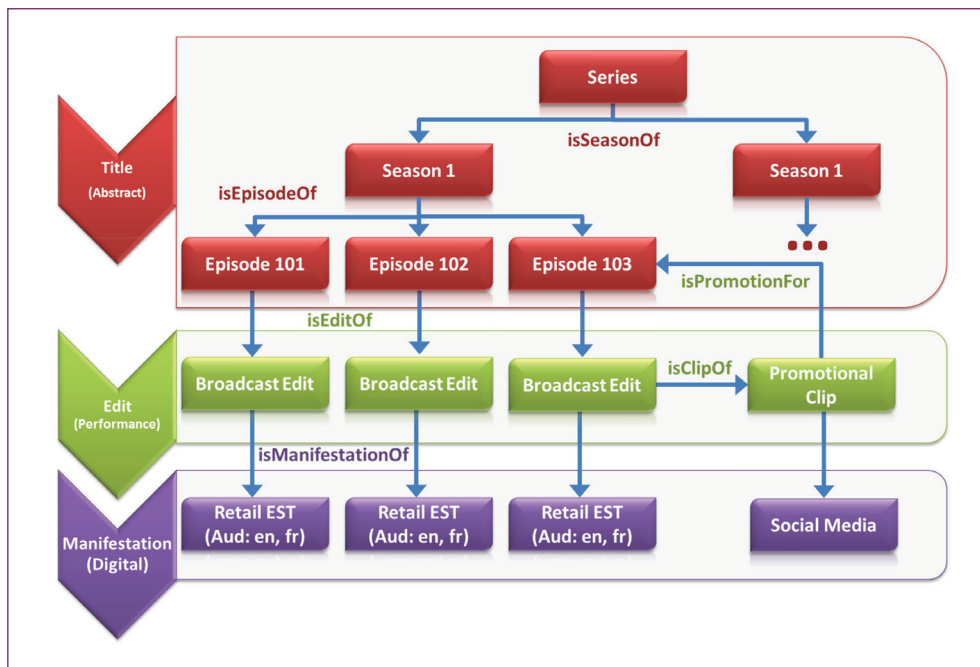
Figure 3. Sample EIDR registration tree for episodic works.

In addition to the hierarchical structure associated with most Content ID records, EIDR also supports a number of different non-hierarchical relationships, such as the isPromotionOf relationship that can link a trailer to the work it promotes. EIDR also supports different grouping relationships, such as isCompilationOf that can be used to identify collections of works ranging from the oeuvre of an auteur director to the contents of a Blu-ray disc. If a particular item or relationship is not already in the EIDR Registry, any EIDR member[53] can create it.[54]

The EIDR Registry strictly controls modification access to existing records, but freely allows Members to create new root records or children of existing records. That way, if an archive ever needs an EIDR ID for a particular application, it is always readily available. Furthermore, this model allows trusted sources, such as producers, metadata providers, and archives to improve the EIDR metadata — adding a missing director, adding a new alternate ID, correcting a foreign-language title, refining a release date, and so on. This can be thought of as a controlled version of crowdsourcing, combining the benefits of group input with a healthy dose of quality control.

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.[55]

Archives will continue to use their own identification system for strictly internal operations, but can gain significant advantage by using EIDR IDs for their interactions with the outside world, and with each other.

---

53 EIDR Members span the media and entertainment ecosystem including producers, distributors, broadcasters, archives, and data aggregators in the Americas, Europe, and Asia. ("Members," http://eidr.org/membership/)
54 EIDR, "Registry User's Guide."
55 "Semantic Web," http://www.w3.org/standards/semanticweb.

4.   EIDR-Enabled Applications for Archives

To appreciate how the broad adoption of EIDR Content IDs can benefit archives, we present a number of scenarios that illustrate how EIDR IDs can be used to enable or improve workflows.

4.1   Federated Search and Discovery

A unified mechanism for search and discovery across archives allows individual archives to act as nodes in a larger archive network, providing transparent client access to the aggregate catalog.

If archives were to integrate curated content IDs into their technical infrastructure, they could more easily provide their clients with the ability to discover, curate, transfer, loan, and provide access to digital content. Such a system could also be extended to search and discovery for physical assets, though transfers and loans would still be more difficult to implement.

Since internal identifiers are rarely shared among archives, a common identification practice in current use within library catalogs is the concept of preferred titles or uniform titles to identify works and their various versions and manifestations. This facilitates manual catalog matching, but is not particularly useful in automation.

Returning to Blade Runner,[56] using EIDR Content IDs instead of title strings, the various incarnations of the film are unambiguously identified as:[57]

- Root Record
  - 10.5240/EA73-79D7-1B2B-B378-3A73-M — the abstract title (for a general search)[58]
- Child Edit Records
  - 10.5240/6E98-7A29-CD01-DFA3-4218-H — the original 1982 theatrical release
  - 10.5240/7290-C8AD-12BA-4F93-3B07-7 — the 1992 director's cut
  - 10.5240/E591-87E8-122F-F5F5-FEF8-3 — the 2007 "final" cut

It would then be a trivial matter to check which archives held copies of the various versions of *Blade Runner* by querying against the desired EIDR ID. EIDR manifestation IDs (child records of a particular edit) could then be used to distinguish a 35mm film print from standard definition video and high definition video, allowing researchers to query and compare not only the versions of titles held at different archives, but also to identify which archive has the work in the optimal format for each particular use.

An archive wishing to find more information about a work, for example for a special screening or to further its own research goals, is able to use EIDR alternate IDs to communicate with the issuing archive (or other sources) even if those sources do not expose EIDR in their public search systems. That way, merely registering a work with EIDR and providing a local identifier increases the possibilities for collaboration even if a fully federated search is not possible.

4.2   Acquisition of Off-Air Recordings

Most national, local, and commercial television and radio archives record off-air television and radio recordings 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

---

56   As the Director did so often.
57   EIDR Content IDs also can be assigned to non-linear works, such as video games, if required.
58   All EIDR Content IDs can be resolved on-line via the EIDR registry and the DOI proxy. In this case, the links would be https://ui.eidr.org/view/content?id=10.5240/EA73-79D7-1B2B-B378-3A73-M and http://doi.org/10.5240/EA73-79D7-1B2B-B378-3A73-M, respectively.

every economy of scale must be applied to make this acquisition activity practical. Archives can collect these materials in digital form using automatable, scalable, and mature off-air capture platforms, such as Cambridge Imaging Systems' Orbital,[59] but the captured descriptive meta-data 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. There may actually be multiple matching exercises per program if the metadata are obtained from multiple sources.

Using EIDR Content IDs for program identification, and then embedding those IDs in the broadcast stream, EPG data, certified broadcast logs, and descriptive metadata available from the broadcasters themselves and third party data aggregators, would allow the archives to collate these disparate sources in an automated workflow to produce a rich content record for down-stream catalog activities, including client search and discovery. This would also allow sharing among archives, so if one possesses an improved metadata set, others can take advantage of it without having to absorb the full cost and time delay to collect and curate the data on their own.

This process would also add value to the recordings themselves, making them more available to the original broadcaster or other licensees for reuse and resale across the broadcast sector. Archives could also make use of EIDR IDs to describe and identify any derivative versions of the broadcast programs they may produce in-house, such as an edit where they have pulled the blacks (removed the commercial breaks) or a package of contiguous clips where they have divided a longer program into more easily digestible chunks. Finally, multilingual programs presented in multiple markets and identified by a suitable EIDR edit or manifestation ID could be automatically linked together via their common abstract title ID by traversing the EIDR hierarchy from the child ID to the parent ID.

4.3   Bulk Digitization

Many archives are moving to digital encodings of their holdings. This offers many advantages, including avoiding wear on physical assets and the ability to deliver content to more people in more places and more ways than ever possible with physical assets and analog media. Before these advantages can be realized, the assets must be digitized. This poses its own challenges when an archive is digitizing its own holdings, but there is an extra layer of manual effort and opportunity for error when third parties are involved.

For example, for the British Film Institute's Unlocking Film Heritage project, the BFI is digitizing 5,000 films from its own collection and acquiring digital versions for an additional 5,000 films from partner archives and commercial rights holders across the UK. For the acquired assets, the supplying partner must find the title in the BFI catalog, correctly associate it with the asset they are digitizing, and deliver the final digital copy along with any metadata enhancements to the BFI for ingest into the BFI's digital catalog. This involves multiple manual touch points, each of which adds to the partner's cost and increases the systemic chance for error. Taking an actual case, there are six moving image adaptations of the play *Hindle Wakes* in the BFI catalog, each with essentially the same description, since they are all based on the same underlying material. During the selection workflow, one of the versions was matched to the wrong BFI catalog entry for both the digital asset and the incoming descriptive metadata, requiring significant time to correct the error. Because of the potential for this sort of mismatch, the BFI must perform manual quality checks on each incoming asset to make sure it has been identified and ingested correctly, increasing the BFI's internal costs for every asset acquired in addition to the cost of correcting any errors that are uncovered.

---

59 "Orbital Enterprise Video Recording," http://www.cambridgeimaging.com/orbital_enterprise_video_recording.

If the programs and their associated metadata were identified with EIDR IDs, most of the manual touch-points could be removed from this process, making it faster, more accurate, and less costly for all parties involved.

In addition, there are cases where the BFI does not have a record for the incoming work. The current process calls for the supplying partner to provide descriptive metadata using a template. The BFI then manually extracts the necessary descriptive metadata to create their catalog record. If the work came in with an EIDR ID, the BFI could take advantage of the alternate IDs associated with many of the EIDR records to automate the acquisition of descriptive metadata from third parties such as IMDb but also including Netflix, Flixster/Rotten Tomatoes, and Internet Video Archive. Internal IDs unique to a work's producer or distributors stored as EIDR alternate identifiers, such as those from 20th Century Fox, ITV, Sony Pictures, or Warner Bros., can be used to obtain descriptive metadata directly from the work's creator.

4.4   Intellectual Property Rights

The EIDR Registry does not store intellectual property rights information or make any assertions regarding rights or ownership. Its sole focus is accurate identification, with support for discovery and de-duplication as necessary companions to achieve that aim. Any system in which rights holders can assert their rights in different works and then provide a mechanism where those rights assertions can be aligned with materials held by archives must first begin with reliable identification, de-duplication, and discovery. Rights certainly depend on factors external to the work itself, e.g., media or distribution channel, territory, and dates, but also on the particular version of a work, covering different music, voice talent for dubbing, edits that add or remove footage controlled by third parties, and so on. The EIDR hierarchy covers all of these internal aspects, allowing the external legal and commercial rights to be connected to exactly the right version of the work.

The Copyright Hub is a UK-based initiative to create the technological infrastructure to support just such a rights tracking system.[60] Identification via curated EIDR IDs is critical to the pilots currently under development, since this will allow the creation of the machine workflows necessary for the process to scale. The assurances of EIDR-based identification build confidence in the process and facilitate negotiation of the delicate rights relationships between archives and their donors. The archives do not want to risk jeopardizing their donor relationships with rights holders by exploiting their content inappropriately. At the same time, archives want to be able to explore innovative content exploitation initiatives without having to allocate manual labor necessary to identify and enforce these rights for every asset held by the archive.

Similar problems exist with orphan works and the diligent search initiatives that are being developed to allow interested parties to legally and safely exploit them, including OHIM and FORWARD in Europe[61],[62] and OWLS in the UK.[63] Such schemes encourage the use of permanent, external, unique identifiers, such as EIDR IDs. The effort necessary on the part of an archive to achieve diligent search certification can be amortized across multiple services if the archive submissions and the corresponding entries in the orphan works databases are identified by an EIDR Content ID. In addition, the alternate IDs in EIDR can provide other sources to search.

Finally, the machine automation facilitated by EIDR IDs would also make it easier for the orphan work services to correlate their data with third parties, such as retransmission rights collectives, to better identify the owners of assumed orphan works (and to more certainly assert that particular works are, in fact, orphans).

---

60   "Welcome to the Copyright Hub," http://www.copyrighthub.co.uk.
61   "Orphan Works Database," https://oami.europa.eu/ohimportal/en/web/observatory/orphan-works-database.
62   "Framework for a EU-wide Audiovisual Orphan Works Registry," http://project-forward.eu.
63   "Copyright: UK orphan works licensing scheme," https://www.gov.uk/government/consultations/copyright-uk-orphan-works-licensing-scheme.

5.   Summary

A few moments spent with a blank whiteboard will lead to the identification of many more use cases where a curated identifier (globally unique, permanent, and universally accessible) can help reduce costs, increase accuracy, facilitate new services, and make an archive's collections even more accessible and valuable to their service communities. The value of curated identification grows significantly with each additional party participating in the ID ecosystem, with each additional database and workflow that carries the IDs, and with each additional asset so identified. Individual archives can also take advantage of the network effect originally described by Metcalfe's Law, where the value of a network, and therefore the value apportioned to each participant, is far greater than the sum of its individual parts. To achieve these goals, a curated identification system must have a global reach and accurately and unambiguously identify the various types of works held by audiovisual archives, from the abstract title level through their derived edits down to their specific representations with clips, composites, and compilations alongside. The Entertainment Identifier Registry Association provides just such an identifier in the form of its EIDR Content ID, based on the ISO-standard DOI system.

6.   Bibliography

Agranoff, David, Walter Michel, and Todd Wakai, "Streamlined Content Metadata Integration and Management Using Entertainment ID Registry (EIDR)." Accessed November 27, 2014. http://www.eidr.org/documents/Expo2012_Technical-PaperEIDR.pdf.

Briscoe, Bob, Andrew Odlyzko, and Benjamin Tilly, "Metcalfe's Law is Wrong." Accessed November 9, 2014. http://spectrum.ieee.org/computing/networks/metcalfes-law-is-wrong.

British Film Institute, "Advanced search." Accessed November 16, 2014. http://collections-search.bfi.org.uk/web/search/advanced.

Business Dictionary, "Metcalfe's Law." Accessed November 9, 2014. http://www.businessdictionary.com/definition/Metcalfe-s-Law.html.

Cambridge Imaging Systems, "Orbital Enterprise Video Recording." Accessed November 24, 2014. http://www.cambridgeimaging.com/orbital_enterprise_video_recording.

Copyright Hub, "Welcome to the Copyright Hub." Accessed November 24, 2014. http://www.copyrighthub.co.uk.

Drewry, Raymond and Don Dulchinos, "Transforming Entertainment Through Technology," M&E Journal (Winter 2013-14): 81-86.

Entertainment Identifier Registry Association, "EIDR System version 2.0 Registry User's Guide." Last modified August 13, 2013. http://eidr.org/documents/EIDR_2.0_Registry_User_Guide.pdf.

Entertainment Identifier Registry Association, "Overview." Last modified July 2014. http://eidr.org/documents/2014-07-01-EIDR_Overview.pdf.

Entertainment Identifier Registry Association, "About EIDR." Accessed November 16, 2014. http://eidr.org/about-us.

Entertainment Identifier Registry Association, "Members." Accessed November 27, 2014. http://eidr.org/membership.

EU Observatory, "Orphan Works Database." Accessed November 24, 2014. https://oami.europa.eu/ohimportal/en/web/observatory/orphan-works-database.

Forward, "Framework for a EU-wide Audiovisual Orphan Works Registry." Accessed November 24, 2014. http://project-forward.eu.

Intellectual Property Office and Department for Business, Innovation & Skills, "Copyright: UK orphan works licensing scheme." Last modified July 4, 2014. https://www.gov.uk/government/consultations/copyright-uk-orphan-works-licensing-scheme.

International DOI Foundation, "Key Facts on Digital Object Identifier System." Last modified October 13, 2014. http://www.doi.org/factsheets/DOIKeyFacts.html.

Internet Movie Database, "Blade Runner (1982) Release Info." Accessed November 27, 2014. http://www.imdb.com/title/tt0083658/releaseinfo?ref_=tt_dt_dt#akas.

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." Last modified October 2012. http://www.itu.int/rec/T-REC-X.667-201210-I/en.

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

Library of Congress, "Library of Congress Classification." Accessed November 16, 2014. http://www.loc.gov/catdir/cpso/lcc.html.

Wallaskovits, Nadja and Christian Liebl, "UMID – Unique Material Identifier." Accessed November 16, 2014. http://www.digitalpreservationeurope.eu/publications/briefs/UMID_Unique%20Material%20Identifier.pdf.

Wittgenstein, Ludwig, *Tractatus Logico-Philosophicus* (New York: Harcourt, Brace & Company, Inc., 1922), Proposition 7.

Wooley, Sheb, "The Purple People Eater" (New York: MGM, 1958).

World Wide Web Consortium, "Semantic Web." Accessed November 27, 2014. http://www.w3.org/standards/semanticweb.