Getting to ROI: Rights metadata and the smart content life cycle

Received (in revised form):

Edward Klaris
is an investment banker, management consultant and attorney, specialising in IP companies and the technology that enables content monetisation. Ed started his own firm, KlarisIP, after having been in the media and entertainment industry for more than 20 years. He is also Managing Director at Desilva & Phillips, a leading boutique investment bank, where he focuses on digital media and e-commerce acquisitions and sales.

Add

Abstract Calculating ROI on individual content assets has long been a dream of content companies. Even the most sophisticated media and entertainment companies cannot be exact about how much a piece of content has made in the past, and what its lifetime value might be. However, some companies are doing a lot to create rich data about their content assets by investing in systems that track the content usage, and by creating metadata not only to describe the content, but also to provide rights, financial and usage data that can be linked to the binary content asset (ie the photographs, video or music). Storing rich metadata in databases rather than in Excel or offline will help content companies to know, protect, value and monetise their content efficiently and broadly. Predictive analytics, machine readable licensing, and ultimately complete content ROI will all become more important to media and entertainment companies competing in the complex digital ecosystem where optimising channels, territories and platforms is becoming more crucial. Transparency into what content performs better than other content, and how much to invest in content will not only lead to short-term monetization opportunities, it will also lead to intellectual property value creation for content when companies look to sell and need to point to the quality of thier metadata to prove the worth of the company overall.

KEYWORDS: Intellectual property, IP valuation, metadata, monetisation, return on investment, rights, technology, technology enabled content

INTRODUCTION

The content business has become a lot more complicated in the past few years. Not long ago, a photograph published in *Vogue* was used in print once and that was it. That photograph may have cost US$100,000 to produce — flying a famous photographer to Europe first class, paying US$60,000 for a day rate and engaging assistants and imaging specialists. The only return on that investment was the inclusion of a few photographs in a single edition of the magazine, which was filled with glossy ads. The photographs would then be returned to the photographer to do with as he/she chose. In reality, there was no return on investment (ROI). The business brought in profit and that was good enough. The same was the case for a television show. The show would air and be monetised by advertising. No one could track its value

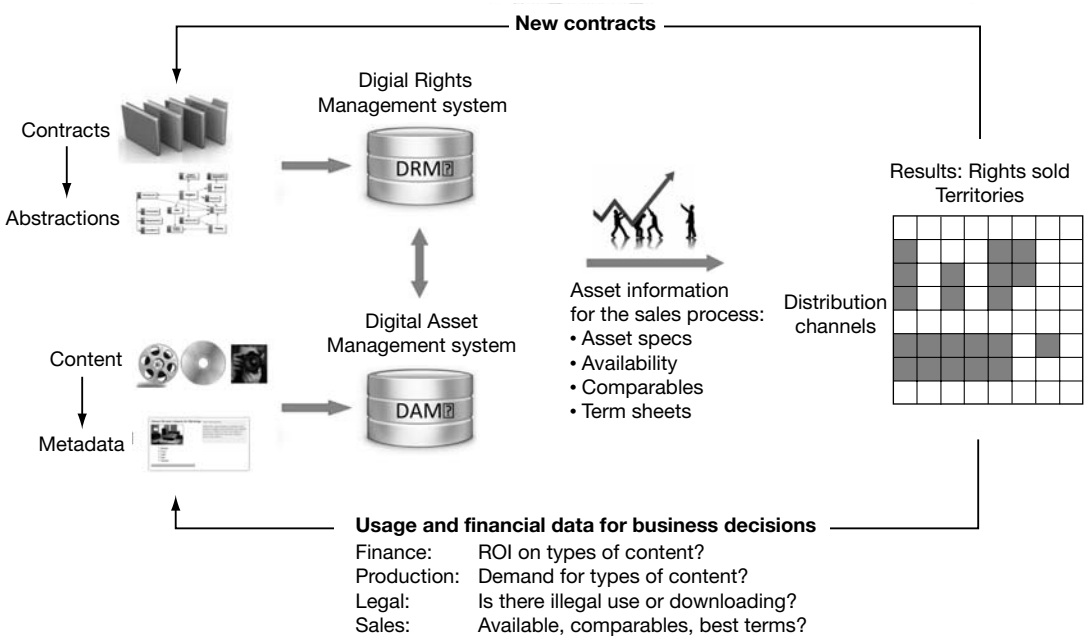

Figure 1: Best practices: The smart content life cycle

and, mostly after just one airing, it would be sent to the network library to gather dust.

What a changed world it is today. There are now many new distribution points: Netflix, Amazon, Google Play, Yahoo, Microsoft, Apple, Hulu, Getty, Corbis, Comcast VOD and countless others will take photographs, scripted and unscripted television shows, movies, music, papers and podcasts from content providers and use them repeatedly for their audiences at a price. Content companies also offer their content direct to consumers through over-the-top and pay walls. Content created in the USA can be delivered all over the world, through many channels and on countless devices. Windowing, syndication, licensing and digital devices have all made it possible to monetise content much more broadly and profitably.

THE VALUE OF METADATA

Calculating ROI in content in this huge and complex ecosystem is a multi-faceted affair. Regardless of what kind of content or what the distribution model is, everything starts — and ends — with metadata. Simply stated, the better a company's metadata, the more asset level information can be tracked, reported on and used to understand the asset's true costs, value and future usefulness. To be clear, metadata are defined in this context as the fielded data associated with a particular binary asset. If the binary asset is an episode of a television show, the metadata will certainly include such items as the director's name, the actors' names and what the episode is about. The metadata may also include whether the show has been licensed exclusively to the UK, for how long and, if possible, what payments were made for the licence of that particular asset.

The metadata for an asset need not all reside in the same database. Oftentimes, digital asset management (DAM) or media asset management (MAM) systems retain the descriptive metadata, making it easy to find, transcode and distribute the asset to

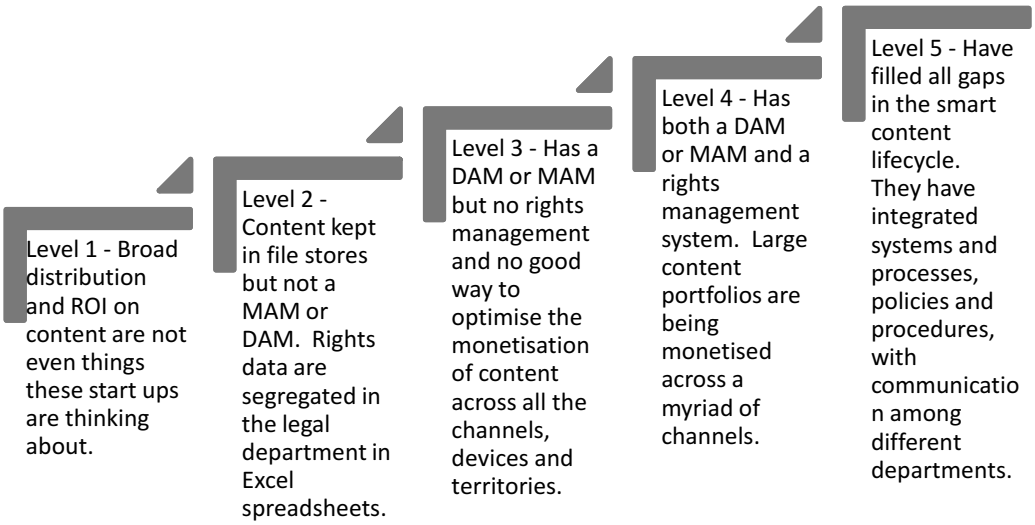

Figure 2: Content monetisation maturity model

third parties or devices. A rights management system will be separate from the asset itself but will reference the asset through a unique identifier and/or an application programming interface (API). Rights data, such as contract terms, financial results and usage data, are all captured in the separate rights database. Still other systems such as the integrated revenue and customer management systems will have metadata referenced to a given asset that controls and tracks information such as subscriber payments and billing.

Metadata can help a number of people in an organisation, from different departments with different functions and needs. Taken as whole, and across various data bases, the systems need to communicate not just among themselves but also to the enterprise resource planning systems (ERPs), where company accounting often takes place. This is easier said than done. Even with the advent of software as a service (SaaS) solutions and open-architecture software design, the process of integrating systems, capturing adequate data and creating reports for all parts of an organisation is quite difficult to achieve. The further along that metadata are found in this 'value chain' of enterprise integration, the more sophisticated discussion there can be about the 'value' of metadata.

This idea of viewing metadata holistically leads to the central focus. This paper describes what it takes to make content 'smart' through the creation of robust metadata and use of systems that enable efficient content distribution, protection of rights holders and calculation of asset ROI. This is called the 'smart content life cycle'. Although costly and time-consuming, companies with large content portfolios that they would like to monetise across the ecosystem must invest to get their proverbial (digital) houses in order. There is no way to calculate ROI or effectively monetise content without an initial investment in the systems and processes to make those calculations.

A smart content life cycle view greatly broadens central assertions about ROI (Figure 1). With enterprise class integration and a view of metadata across

descriptive, rights, reporting and ERP systems, there can be a conversation about depreciating the costs of metadata. The idea can be set forth that a digital object has appreciating or depreciating value and has book or net present value. It can also be asserted that a smart content object has measurably deepened value for records management, e-discovery and other aspects of risk mitigation.

THE METADATA MATURITY MODEL

Different companies are at different stages of maturity in terms of creating smart content through metadata and systems (Figure 2).

A level 1 company will probably not be thinking about anything in this paper because the company will be focused on its beachhead market with a single buyer/end user in mind. Broad distribution and ROI on content are not even things these start-ups are thinking about.

A level 2 company is one where the initial business is established. Company assets are kept in well-ordered file stores but not a MAM or DAM system, while the only rights data are segregated in the legal department in Excel spreadsheets. The important thing for these companies is to put their content data in a database rather than in spreadsheets. Even if the database just contains the contracts and usage information, once the company gets more mature, it will at least be able to port its simple database fields to a more sophisticated system. Such as Access or MySQL.

A level 3 company has a DAM system but no rights management and no good way to optimise the monetisation of content across all the channels, devices and territories. This company is mature enough to know what it is missing but having trouble allocating resources to the problem. More people keep getting added but systems have not been installed to help monetise content.

A level 4 company has both a DAM and a rights management system. The business is robust and large content portfolios are being monetised across a myriad of channels. Most level 4 companies get all their rights-in contracts data into the system but not their rights-out data. Whatever rights information is in the system was created in-house, without using industry-standard rights expression languages. Reports and tracking are possible but could be better.

A level 5 company is hard to come by. They have filled all gaps in the smart content life cycle. They have integrated systems and processes, policies and procedures, with communication among different departments so that reports and business intelligence are useful and robust for everyone involved. This paper focuses on companies in levels 3 and 4.

RIGHTS AND USAGE METADATA

Most large content companies have become fairly sophisticated about descriptive metadata and asset management. They have DAM or MAM systems, library scientists creating taxonomies and processes for transcoding and distributing files. At Condé Nast, for example, people can find photographs by theme, colour, style, garments, accessories, ingredients and unique identifiers. The asset metadata are all-important and well controlled.

On the rights side of the company, things tend to be quite different. The Office of General Counsel (OGC) resides in its own separate silo from the asset management teams. The lawyers and paralegals in the OGC are not trained in creating taxonomies, rights trees or controlled vocabularies. Data integrity is not part of their DNA. Instead, lawyers tend to look at 'the deal' and to treat each

deal as unique. The words chosen in contracts vary by attorney and by negotiation. Film and television commission agreements can be 15 pages, single-spaced, with long paragraphs detailing every possible eventuality and payment. Publishing contracts can run to many dozens of pages, with selling area definitions hard to discover or captured in odd geographical terms.

Therefore, what does it mean to create metadata from contracts? First, a contract must be abstracted into fields. If, for example, a content licensing agreement is being abstracted, one field might say 'internet'. The box could say 'yes' or 'no' or 'with duties'. If 'with duties' is clicked, then there would be a number of questions to be answered: 'royalties'; 'no cropping or editing'; 'music payments'.

The contract abstracts can take a long time to create. Television and cable companies might have 3,000 contracts a year to abstract. Magazine companies process about 8,000 contracts a year.

To the extent that a contract contains boilerplate paragraphs, companies tend to create a model abstract so that it is pre-populated with rights data every time that boilerplate paragraph is used. Some contracts are templated so that the entire contract can be pre-abstracted. In that case, almost no work needs to go into the abstracting process. When a boilerplate contract is signed without any changes, that contract can be put into a database quite easily. Most media and entertainment companies follow the 80/20 rule with regard to boilerplate contracts. Eighty per cent of all contracts are boilerplate and unchanged. Twenty per cent are heavily negotiated with the other contracting party. Negotiations occur when the other party has leverage and can demand changes. Going off template also occurs when a deal is unique and does not fit the boilerplate.

As companies tend to define legal terms of art uniquely, the abstracts that they develop are also unique. For example, companies tend to define the term 'syndicate' differently. To some, syndicate may mean a distribution right. To others it is the right to license content without naming the original publication. While library scientists rely on specifications such as Dublin Core to provide structure to their metadata, OGC tends to create rights data in a vacuum without regard to industry standards.

In a related development, the World Wide Web Consortium (W3C) is working with the International Press Telecommunications Council to develop a standard rights metadata expression language called online digital rights language (ODRL). ODRL is of interest to news organisations but has not yet been adopted as a true industry standard. Companies other than news organisations can also benefit from standardised rights expression. Content monetisation relies on rights metadata and, as explored below, standard rights expression can have significant positive results.

Some companies think standardised rights expression languages will never be the norm. The issues are many and they are complex. Even if a company is a level 4 or 5, with a rights management system in place, the company will have chosen a system that employs a custom rights tree, not one based on industry norms. Getting the rights software companies to adopt a single rights expression language will be quite difficult. In addition, OGCs tend not to be disciplined about terms and do not see the value of using words across the industry uniformly. In fact, some mature companies consider rights expression to be a competitive advantage and do not want to share how metadata from contracts are expressed.

By contrast, there are two prominent reasons that standardised rights expression languages make sense. First, licensing

content currently requires a human being to be in the middle of the transaction, checking and clearing the content. With standard rights expression, two companies could talk through their machines. One company system would connect to the other system through an API and content could be selected and paid for through a machine-readable contract and avoid the labour costs. Secondly, with standard rights expression, a company could port its rights data from one system to the next. As of now, once a company chooses a rights management system and uploads all its many contract abstracts, the process of porting to a new system seems near impossible, as it probably is. With standard data, porting from one system to the next is easier and less expensive.

Rights metadata involves both rights-in and rights-out agreements. At Condé Nast, a rights-in agreement was the deal between the company, as licensee, and a photographer, as licensor. A rights-out agreement is where the company is the licensor and a place like Getty is licensee. Taking this example further, the Getty agreement must be abstracted and linked to the binary asset through a reference identifier. When Getty licenses the photographs to its end users, a payment is due to Condé Nast. Companies like Getty tend to provide sales data to third parties though Excel spreadsheets. However, Excel is not a robust or efficient way to get the data. It would be far more efficient to have the sales data delivered back to the licensor through an API — and thus deepen the smart content knowledge and the ability to assess ROI. Sales data are metadata. They represent the first bit of ROI data; thus, imagine numerous third-party (rights-out) agreements and sales reports coming back to the content company and being housed in a robust rights management system. Those data can be used to calculate revenue earned on a given asset or ROI.

RIGHTS MANAGEMENT SYSTEMS

Content that is rich with rights metadata is only 'smart' if it is housed in a robust rights management system that can show the content availability and calculate the content's value. This paper will not go into all the different systems available on the market. Some rights systems are installed locally, others are hosted and still others are multi-tenant SaaS platforms.

Rights management systems need to inter-operate with DAM and MAM systems; as with all contemporary application systems, they need to have simple user interfaces, with dashboards for reports and show what content is available for use based on rights-in restrictions and rights-out exclusivity deals. The systems should be able to integrate with ERPs, have robust databases and make it easy to add data from contracts through an easy user interface. The more clients using a rights management system the better because in this quickly changing environment, client feedback for system improvements is crucial to staying current.

At the time of writing, there is no single rights management software that is configured to use a standard rights expression language such as ODRL or MPEG-21. They are all custom expression languages. Some companies are paying attention to the rights standards bodies and may move in that direction soon. This question is essential when examining the future roadmap of a potential rights vendor.

BUSINESS INTELLIGENCE AND FINANCIAL REPORTING

Once a content company has a robust DAM and rights management system in place, the business intelligence and financial reporting needs will follow. The cost to acquire the content is the first piece of financial data needed. This information can

be obtained from a project management system or the ERP system and from rights-in contracts. The rights management software houses all 'duties' and 'constraints', which set forth all the royalty or flat fee payments that a content company must make if a given piece of content is used in a particular way.

Some rights management software feature standard 'out of the box' reports and will likely feature custom reporting capabilities. Many times, larger companies can engage a financial analytics firm to build all the data queries needed in systems like Cognos by IBM, which use multi-dimensional and predictive analytics tools. The reports all become part of the metadata associated with the assets.

Revenue data can be obtained from multiple sources, such as advertising and circulation revenue, subscriber, licensing royalties, affiliate revenue and retransmission fees. With this information, a content company can begin to ask more sophisticated questions, such as:

- 'How can I leverage my content to optimise revenue?'
- 'How can I negotiate the best deals to optimise revenue?'
- 'How can we leverage our own data to predict best outcomes and drive decisions?'
- 'How can we leverage social media to predict market demand?'
- 'How can we optimise our processes to minimise risk?'

Sales people, trying to distribute content broadly, must have availability reports that are updated upon the closing of every rights-out deal. Rights management systems have those reports. Smart content also prominently shows 'where used' and ranks content objects by popularity and, in an optimal setting, correlates the search returns information with demographics and the company's customer relationship management system(s).

CONCLUSION

Robust metadata, which are continually created and associated with an asset, are the key to making content smart. Being disciplined about rights expression languages, financial data and reporting are all important to building revenue and predictive analytics models.

The necessary investment to connect the systems and processes to optimise the content distribution process is considerable. Content companies have to get ever more serious about these issues if they want to stay ahead of the increasingly complicated distribution demands for content.

ACKNOWLEDGMENT

Many thanks to David Lipsey and Howard Weinberg for their help in thinking through and editing this paper.