5 Things Corporate Litigation Professionals Can Learn From Their Records Management & Information Governance (RMIG) Counterparts

By Sandra Serkes
President & CEO of Valora Technologies, Inc.

Introduction

This Valora White Paper examines the relationship between the ways Corporate Litigation groups and Records Management & Information Governance professionals organize, manage and control document data in their respective domains. This White Paper is Part II of this series, and is the companion paper to Part I: 5 Things RMIG Professionals Can Learn From Their Litigation Counterparts.

While Litigation and Records Management & Information Governance (RMIG) departments may have different goals, there are commonalities for how best to manage complex and evolving document populations. RMIG professionals have been curating voluminous information longer than their Litigation counterparts. They have developed certain disciplines, best practices, and lessons learned that can save Litigation "newbies" considerable, time, pain and expense. After too many years of solving document management and analysis problems with ever-more bodies and expense, Corporate Litigation teams are learning to take strategic control of their matter management. Here, then, are Valora's Top 5 techniques for corporate and outside counsel litigation professionals to learn from their RMIG counterparts.

1.0 Data is Not Always a Liability. To RMIG, Data is an Asset.

Is stored information, what we at Valora call "DDC" for Data, Documents & Content, an asset or a liability? The answer depends, of course, on how well you manage and control your DDC. If you control them well, that is to say you *curate*[1] them, then DDC are an asset. The asset view is a common refrain amongst RMIG professionals. In their "Big Data" view, DDC help identify issues and trends, revealing important business insights from otherwise un-helpful stored information.

---

[1] The Council on Library & Information Resources defines data curation as "the active and ongoing management of data through its life cycle of interest and usefulness to scholarship, science, and education. Data curation activities enable data discovery and retrieval, maintain its

However, it is precisely the easy ability to "mine" DDC by identifying patterns and exposing issues that concerns litigators, compliance officers and regulators. To them, DDC are a liability. If DDC is not well managed and curated, then they present an unknown, and thus highly dangerous, liability.

Interestingly, both sides agree on one critical fact: the better managed the DDC, the safer/more informed everyone in the organization becomes. Thus, the trick is to get strategic control of the DDC and remove or mitigate the dangers, while harvesting the benefits. For litigation professionals, this means moving *beyond* the common ideas of:

a) Culling as the only form of DDC control
b) Looking at one litigation matter in isolation from other litigation matters and/or other corporate uses of information
c) Collection & production as the only mechanisms for the flow of DDC in and out of the organization

Culling is not a DDC organization strategy

Litigation professionals often cull document populations to get to the most relevant, the most damaging, or the most vindicating material. To do so, they remove irrelevant material from the database or collection, a process known as **culling**. To be fair, litigation teams often find themselves weeding through a lot of junk: irrelevant emails, spam, announcement blasts and other miscellany. The desire to cull, cull, cull is strong and well-ingrained. At its best, culling reduces the overall population so there is less to actually sit down and pore over (aka Document Review), there is less to potentially produce and there is typically less expense to process, host, etc.

While culling surely has its place, it is <u>not</u> a proxy for intelligent DDC management. In fact, pure culling based on relevance or privilege may actually be a mistake to the larger picture beyond litigation. Culling removes documents, and thus information, from the system. A better option than culling (which typically implies removal of DDC altogether), is **suppression**. Document or Content Suppression, sometimes called "quarantining," removes the DDC from active view or use in searches, document review, productions, etc., but not necessarily from analytics or other data mining activities. It also leaves the "culled" DDC from one particular litigation matter available for other matters or other uses. We'll come back to the notion of DDC re-use in Section 3. For now, let's simply highlight the idea that intelligent management of information, oriented around suitability for a specific purpose, is a much smarter strategy than wholesale culling.

Document Populations are Not Static – Even in Litigation

In Litigation, as in many business processes, additional documents and content frequently enter the picture after the initial set is collected. Sometimes new custodians' data enter the mix, sometimes material comes from third parties or FOIA requests. Litigation professionals can learn from their RMIG counterparts about the notions of **Intake, Retention and Disposal** of documents, as part of an ongoing, expected process, rather than as an emergency fire drill each time it happens. (Evolving from perpetual fire drills to rational process is a theme we will

---

quality, add value, and provide for reuse over time, and this new field includes authentication, archiving, management, preservation, retrieval, and representation."

explore further in Section 5.) Most RMIG teams have mapped out a regular workflow[2] of documents into, through, and out of their document management systems.

Intake can coincide with collection efforts, and often includes some kind of data processing and/or document massaging as part of the ingestion into the ultimate database or hosted system. For example, Intake might include a special email Inbox or an online intake form where custodians upload relevant DDC. Larger DDC collections arrive via ftp or on hard drives, often with a first stop being some kind of file processing (usually for text and metadata extraction, and/or duplicate detection and removal). By combining *arrival* of documents with *preparation* of the data for analysis, hosting and review, smart teams reduce steps and save time and money. In doing so, the set up an ongoing, repeatable document workflow for additional documents in a single matter, or for additional matters and purposes.

Just like corporate records management, litigation teams should establish rules for document retention. The simplest rule, already being followed in most cases, is **Relevance**. Relevance asks: is the document's content, attributes or source in some way germane to the matter in question, to other matters or to the client in general? Non-relevant documents are culled (or perhaps suppressed, per above), and not retained in an active sense. There can be many other nuances to retention, such as date windows or the presence of certain terms or attributes of the DDC. Following RMIG models, litigation teams should undertake actual documentation of the rules by which documents will be retained, for how long and in what manner, essentially, a **Litigation Matter Retention Schedule**. Once this is in place, it is simply a matter of testing each subsequent DDC item against those rules to determine inclusion, suppression or removal of the item – a task which can and should be automated.

2.0 Data is a Mess!

Smart RMIG teams know all too well the kinds of problems that exist inside collected DDC. Documents are out of order; they're misnamed; they have the wrong metadata; columns, pages or fields are missing; they're in foreign languages; they've got passwords, viruses or broken links; they're riddled with PII[3] or other sensitive company information. The list goes on and on. Litigation teams, however, are often surprised by such messy DDC events. They often falsely assume that incoming DDC are clean, and free from such hazards. This is a costly and painful mistake. From Valora's experience and the vast experience of industry groups like AIIM and ARMA, it is better to assume the worst and perhaps be pleasantly surprised.

Fortunately, these messy data DDC problems are so common that there are numerous ways already in place to help deal with them. With tools and solutions such as AutoTranslation, and AutoRedaction, sensitive or foreign content documents are easily converted and redacted. AutoUnitization and AutoIndexing solutions help organize and tag incoming DDC for easy analysis and categorization. Automated NearDuplicate Detection solutions help classify and organize similar content across documents.

608,087,870

– total number of records containing sensitive personal information involved in security breaches in the United States since January 2005

Source: Privacy Rights Clearinghouse, June 2013

---

2 The following links provide some examples of document workflows. Digital Records Workflow. Paper Records Workflow. Billing Workflow. EHR Workflow.

3 PII stands for Personally Identifiable Information, such as personal addresses, phone numbers, Social Security Numbers and similar.

Litigation teams would be wise to assume and expect that these types of problems will exist with their DDC and build solutions into the Intake & Collection of their materials from the outset. Smart litigators seeking Requests for Production should take the further step to demand integrated solutions for opposing parties' materials, too. For example, it is reasonable to request that opposing parties pre-redact their documents prior to production, such that the burden of identifying and determining sensitive information falls upon the party most knowledgeable about its contents and thus most able to perform such tasks.

RMIG teams have learned not to rely on the creators or storers of DDC to organize and tag materials appropriately. Most people will not fill out complicated metadata fields, or forms, or even use established folder structures for organizing content; it's just too much trouble. Fortunately, those same automated solutions mentioned above obviate the need for users and content creators to mark and tag their work product. Automated solutions will automatically:

- tag and classify content
- move (or copy) documents and files to the correct locations
- organize by importance, relevance, duplication or affiliation
- put appropriate permissions & restrictions on information
- notify interested parties when DDC are available, aging, need attention, etc.

3.0 Understand & Champion DDC Reuse

RMIG groups have long championed the idea of Information reuse. Generally, the theory is that information (an asset) has many purposes and can be used again easily if stored, managed and retrieved properly. However DDC reuse is a new concept for most litigation teams. This is in part due to the "one-off" legacy of litigation. Historically, once a matter was done, that was that. Given that today's litigation teams are typically mining and collecting DDC from active, ongoing information stores, the litigation DDC do *not* particularly live in an isolated context anymore. At best, they are copies of DDC already in use or reuse elsewhere in the client organization. In fact, litigation teams are frequently **duplicating a lot of effort and content**, just to create an artificial litigation document environment.

This strengthens the Return on Investment (ROI) case for organizing & controlling DDC prior to litigation fire drills. In other words, proper DDC management (RMIG-style) helps make litigation DDC needs easier to manage. Instead of each litigation matter having to bear the cost burden of information collection, mining, culling, analysis and hosting, this work should be done once, on a global basis, with periodic updates to stay current. Then, when the next litigation matter arrives, it is simply a matter of selecting the pre-managed content for the appropriate uses. With proper RMIG-style DDC management, any litigation matter can benefit from the efficiency and process that is normally reserved only for the "bet the farm" matters.

> *Smart Data, Document & Content Management should **transcend** any particular litigation*

As a final selling point to RMIG-style litigation DDC management, consider that RMIG groups do not typically worry about having to prove defensibility of their data collection and analysis methods. They are defensible by definition. That is, they are highly repeatable, transparent, and produce the same results regardless of who actually performs the work. For any litigation teams who've had to argue their process defensibility in court, this should be a big selling point.

4.0 The Perfect Data Myth

There is no such thing as perfect information or perfect data. Smart, yet underfunded RMIG teams (which are virtually all of them), understand the reality of cost-benefit tradeoffs. In fact, like messy data, imperfect data is a fact of life in the RMIG world. The general argument is that DDC are basically well-controlled for the most part, and the occasional error or missing information is an "acceptable casualty" within the range of normal operating procedure and manageable cost structures used to maintain such systems.

Litigation owes its legacy, and certainly its notoriety, to large, high-profile, "bet the farm" matters, in which success is aimed for *at all cost*. However, most litigation matters today, and indeed most litigation clients, are *not* in a position to pursue their claims without respect to the cost of those activities. Instead, a rational, thoughtful and aware approach to litigation has been emerging and this includes the DDC management of matters, arguably one of the most obvious areas in which to control costs.

Litigation professionals would be wise to learn one of the basic tenets of RMIG: Build systems for the *most-common* or *most-likely* scenario and then build in, expect, and budget for **exception handling**. Sometimes called the 80-20 rule, the most cost-effective systems will handle at least 80% of the "typical" DDC. Are most of the documents business reports, email correspondence, financial data? Do documents need to be locked down or can they be reviewed remotely on individual devices?

Exception handling lays out an explicit plan for handling exceptions as they arise. What happens to the foreign language document? What about the password-protected file? By laying groundwork for exception handling at the outset, smart RMIG teams a) assume there will be exceptions and b) have an established model for handling them. This keeps down costs, effort and angst – good goals for litigation teams, too.

Another core RMIG principle is: Learn to live within a modest budget. Unlike litigation, RMIG's legacy is one of backroom stepchild having to "make do" with small budgets and minimal executive attention. While this is changing in the era of "Big Data," high profile Litigation and eDiscovery teams would be smart to learn to conduct litigation DDC management modestly as well.

Below are 4 examples of smart, low cost, and efficient DDC management & control.

- Change the view on expenses from "at any cost" to "reasonable expense" and adopt a ROI mentality
- Look beyond single documents. What does the data, the population, the trending tell you? Where is the pattern of risk & exposure, as opposed to what does this one document say?
- In preparing and reviewing documents for a matter, ask what can be done to assist future matters down the line? What processing, information, analysis or dispositions can be utilized again?
- Consider which documents can safely be disposed of due to their obsolescence, irrelevance or duplicativeness.

5.0 Litigation DDC Management & Ownership

In fighting their influence battles, RMIG professionals have learned to make DDC management & mastery someone's explicit responsibility. It is common to find people with the title of Director of Records Management or Chief Compliance Officer. The closest Litigation & eDiscovery title

would be Litigation Support, which encompasses many non-DDC tasks, and typically holds a second class standing to titles with "Attorney" in them (whether contract or otherwise). In litigation matters, who is responsible for the overall intelligence, cost-efficiency, use, reuse and exception handling of the matter DDC? Is it the senior partner responsible for the whole matter? The paralegal responsible to 15 partners? Who will make sure DDC is managed efficiently, with an eye towards the best, most reasonable solution? It could be outside vendor or consultant, but it is far better for it to be someone on the case team, within the corporate legal department, responsible for many/all cases eith this as their *primary* job.

Above are just 5 smart lessons that Litigation and eDiscovery departments can learn from their RMIG counterparts when it comes to large-scale document databases. There are also **numerous lessons you can teach them**, which is the topic of the previous White Paper in this series, the companion to this one: 5 Things Records Management & Information Governance Teams Can Learn from Litigation & eDiscovery. For more information on these topics and many more, please visit Valora's website and blog, or contact us at: 781.229.2265.

About Valora Technologies, Inc.
Valora is a technology-based provider of automated document management, analysis & review services for the legal, records management & information governance industry. We offer data mining, analytics, document intake and visualization, and hosted solutions for corporations & government agencies, as well as their advisory, inside & outside counsel organizations around the world.

Valora has developed a strong expertise in the processing, management & analysis of large and small matters with complex requirements, such as short deadlines, sensitive material & mixed languages. Our specialty is providing efficiency, organization and cost control.

About Sandra Serkes

Sandra Serkes is a dynamic leader with an extensive background spanning 25 years in software marketing, product management & corporate strategy, particularly in document processing & analytics, computer telephony & speech recognition. Today, Ms. Serkes oversees Sales & Marketing, Finance & Administration, Operations, Engineering and Corporate Strategy at Valora Technologies, Inc., where she is CEO, President & Co-Founder.

A graduate of Harvard Business School and MIT, Ms. Serkes is a frequent industry speaker & panelist. She is an active participant in the Women Presidents' Org., The Commonwealth Institute, the MIT Enterprise Forum, the MA Software Council & the Network of Harvard Alumnae. Ms. Serkes serves on the boards of several technology and service start-ups, and was named a 2006 "Woman to Watch" by Womens' Business magazine. Ms. Serkes will be the upcoming Keynote Speaker at the ARMA Northeast Regional Convention in Boston in 2014.