Derek Sturdy

References

[1] http://www.sims.berkeley.edu/research/projects/how-much-info-2003/
[2] StorageTek presentation at Storage World 2003 Conference

Biography

Adrian Dale is renowned as a radical thinker in the information and knowledge management fields - on a mission to integrate the divided camps of the information management world. He has managed the information, records and knowledge functions of several UK organisations in the public and private sector - handling the shift from the paper to electronic worlds. He coaches the information management teams of a number of public and private sector organisations, helping them creating the radical change that is necessary today. Adrian is Editor of the *Journal of Information Science* and Chairman of the Online Conference 2007. He is a fellow of the CILIP and a Visiting Professor of Knowledge Management at the Athens University of Economics and Business. Adrian.Dale@creatifica.com

Enterprise Content Management

Abstract: In this article, which closely relates to Adrian Dale's topic for the Willi Steiner Memorial Lecture, Derek Sturdy explains what ECM is. He discusses the storage, editorial and publishing processes involved in ECM. He explains the importance of workflow and considers the advantages and disadvantages of potential software solutions to manage the process.

Keywords: Enterprise information management; content management; software

Introduction

It's hardly a great line at the Saturday night party: "I'm in Enterprise Content Management". Hmmm. At worst, you think of losers in non-jobs jargoning, with terrible earnestness, as if their lives depended on it. At best, you mournfully expect seminar-speak – here's an example:

> "[Business units] desire applications that let them utilise content more effectively to achieve their line-of-business objectives"

That makes you want to say "well, OK, but they probably don't want to run out of pencils either". (I'm not going to give the attribution of that quotation, to spare the blushes of a senior executive at a major content management software vendor; but I didn't make it up.)

It is the purpose of this article to address the real, and serious, purpose behind the woolly image. It is this. Content is the stuff of communication. The silliest YouTube clip, and the weightiest documentary on tree-shrews in Madagascar, may well share the same format

("media") and even audience: it's the content that matters. Without content, there is only an empty wrapper, useless until it has something to wrap. Managing content has to be a core competence of any organisation more complex than a plumbing business. Since the work product of lawyers overwhelmingly takes the form of the content of documents, content management is critical.

What is Enterprise Content Management?

The "enterprise" piece is just a weasely bit you bolt on nowadays, to make it sound more important.

Content management includes three processes:

- Storage (more or less "records management");

- Editorial (more or less "content capture and creation");

- Publishing (more or less "content distribution").

Many of the functions of these three processes are included, for instance, in document management and

client relationship management systems. Content management is a generic term which is not bothered by the nature of the content. But it also includes four pervasive processes which are applied to all three of the core processes listed above:

- Managing wrappers and formats (mainly a content management function);
- Managing workflow, the units which make up business processes;
- Managing compliance, with both external (regulatory) and internal standards;
- Managing access and security.

The last three of these are not specific to content management but will crop up in many aspects of a legal organisation's work. Because they are general, they are not discussed in detail here; but that does not make them any less part of the content management function. Compliance, in particular, is seen as a major driver behind the whole "records management" initiative (think, for instance, about discovery), as well as influencing how certain sorts of content is formatted and structured for specific tasks.

Content management is deeply connected to access and security, since it is often the content of the materials that defines the actual application of access rights and security provisions. It's a huge topic, but it is one which we can largely relax about, (which does not mean we ignore the issues, or fail to take them seriously). The document management systems commonly in use in legal organisations do a fine job in this area, and the underlying technical security issues are the worry of your IT Director.

Managing storage

Storage is about ensuring that content can be found (or not found, if that is important) by people inside or outside the organisation. The vital pieces are:

- Where the content is to be found;
- What applications are needed to read it/view it/ download it/use it/listen to it/change it/remove it/ archive it, etc;
- Dates, authorship and ownership, rights, purpose;
- Additional metadata ad lib., such as:
  - Validity, approval, reviews;
  - Granular security levels, permissions, access rights;
- Update management
  - Is this content to be frozen, can it be changed, should it be changed as needed?
  - If it is to be changed, the processes and reporting such as link management and testing,

and automatic change and testing where underlying formats have become obsolete, need to be put in place;

- Some dormant content may need to be re-entered into the editorial processes to return it to live status.

Managing content creation and editing

This is familiar stuff and so only a summary is needed here:
Mechanics:

- Templates and rules for creating content;
- structure;
- Formats – appropriate tools for creating complex formats (such as .pdf and XML) from familiar ones (such as Word);
- Link management;

Collaboration – managing multiple authorship, whether internal, external or both;
Control:

- Identifying change, controlling change, recording change;
- Approval processes;
- Version control;
- Access control and security;
- Metadata creation, including classification, and management;

Update schedules (which feed into update management, here associated with storage management as outlined above).

Content capture

This is a cognate set of processes to internal content creation and editing where the origin of the content is not within the organisation:

- Mechanics:
  - Storing hard copy materials (increasingly a secondary method)
  - Scanning and imaging
  - Moving content from emails, whether in the email body or as attachments
- Rights management:
  - Is the content public domain?
  - If not, what are the associated rights, and will they change over time?
  - Do the rights affect the management of the content – eg storage media?

161

- Content integration:
  - Is particular captured content to become part of internal content - effectively treated as collaborative content?
  - If so, internal standards for wrappers and formats (see below) need to be applied to it, and
  - What happens to the "as received" content (which becomes a storage management issue).

Publishing

Content distribution has its own processes:

- Mechanics:
  - Mostly, for LIM readers, text and graphics are all that needs to be published, but there is still the issue of selecting the application which gives users access to the published content;
  - Destination: printed materials, intranets, extranets, web sites, emails;
  - Format translation, based on the destination (eg Word to XML);
  - Automated content integration (eg names and addresses from CRMS to a mailing list or to specific letters);
- Integration and organisation (may be part of the content creation and editing workflows)
  - Pulling content from various sources into a new aggregation (as in bulletins);
  - Ensuring that the published content has context, eg the "real property" section of a website devoted to M&A procedures, or attaching an email to a matter;
  - Adding metadata to the published content, including rights metadata.

Wrappers and formats

Content management includes the nesting of content and the rules by which it is done within the organisation. If content wrapping is allowed to be at the discretion of any content creator or editor, there can be serious implications not just for the mechanics of content storage, but also for compliance issues of many kinds. Examples of content wrapping are:

- Paragraphs, sections, documents;
- Clips, shoots, films;
- Captions, pictures, text blocks, brochures;
- Rows, tables, databases.

Formats are treated similarly. Part of the content management job is to define the standards that the organisation uses

for text, data, pictures, etc and put in place the process to ensure that content complies with the appropriate standard. It is rare that the organisation needs to worry about creating the standards, which are normally public-domain.

Workflow

Making sure that a business process runs effectively and properly every time can be as simple as applying a bit of common-sense or highly complex (think of processing insurance claims as lying somewhere in the middle). Within the strict content management functions, workflow is appropriate particularly in:

- Content creation and capture – ensuring that rights management, author attributions, acknowledgments, editorial review, content review and acceptance, style review and acceptance, and so on, are all performed in their due place, whether initially or during updates and reviews;
- Content storage – ensuring that the right metadata is applied at the right level of detail for the type of material, and that material is stored in the right place with correctly set access rights and security;
- Publishing – ensuring that content submitted for publication again passes through the correct approval and reformatting procedures and has the appropriate access rights and security applied.

Compliance and conformance procedures, whether imposed internally or by external regulation, are most usefully included within the appropriate workflow. It should be emphasised that having a workflow does not have to mean having a software package: a simple procedure sheet can do wonders. Where work is distributed among several people, perhaps in different locations or departments, a workflow application can become essential to avoid making mistakes and as part of a more general risk reduction policy.

We will return to workflow below, in summarising the most appropriate way to consider it within the content management function.

The software position

Small is beautiful

Point solutions – pieces of software which do a restricted, defined job – started the whole document management – records management – client relationship management suite of software applications. Many of the individual offerings got their start because the original software was written to address a specific problem identified by a particular organisation. In the legal marketplace, one could list literally dozens of splendid point solutions – very good at what they do, often highly tailored to legal work.

The problem, of course, is that content management, as described above, then becomes a matter of tying together all the point solutions selected. The point solutions vendors have reacted to this in a number of ways:

- Simple re-labelling – if people need a "knowledge management" component, let's simply re-label one of our products "knowledge management" and hope for the best (document management and search system vendors have done this, for example);

- Encroaching on each other's territory (portals, email systems and document management systems are good examples);

- Adding on functionality, such as more detailed access rights and security management, or building in some other vendor's search engine.

This article suggests that many of these point solutions offer excellent value for money, especially where the initial development and route-to-market costs have long ago been written off. Below, I will argue that in most cases the clever bit is to ignore some of the bolt-ons, and return to the core functions and purposes that the point solutions were developed to address.

Big is beautiful

Most LIM readers will know that the chance of a particular software introduction going wrong is roughly proportional to the size of the software and the number of things it tries to do at once. It is not just the NHS that struggles with big IT projects. Software introductions which attempt to change several methods of working simultaneously are high-risk.

Fortunately, we can mostly relax about this. The big content management systems, deeply integrated into many aspects and types of business process, are not usually sold to the legal market. There are good reasons for this. Large enterprises have far more employees engaged in routine work, as well as the need to marshal the output of thousands of workers (such as research departments for the big pharmaceutical companies). Legal work may be all about creating and managing content, but the main content producers are the high-profile members of the legal organisations – the lawyers themselves – and the sort of content they produce is restricted in format but almost unrestricted in scope. Few law firms could, even if they wanted to, pay hundreds of thousands of pounds for a large-enterprise type, integrated content management system. They need something more flexible, that they can feel they control.

Where is the future?

If, outside the largest legal organisations, "big software" is not likely to provide a low-risk, cost-effective answer, then the trick must be to turn the various point solutions

into a coherent whole. In other words, the individual solutions, with both their varying degrees of overlap, and their varying degrees of gap, need to be welded into an interlocking jigsaw of functional components.

As the analysis of what content management is all about, set out above, shows, there are lots of pieces. It also shows that some of them lead to others, and that there are components such as workflow, security and permissions, and compliance issues, which are quite pervasive. So the answer to turning the various point solutions into the jigsaw is, of course, in general, to use a portal, and in particular, this article suggests, MS SharePoint.

SharePoint 2007 has come a long way from the early manifestations. Six years ago, SharePoint was a solution in search of a problem – very few people knew what it actually did, which wasn't much anyway. Today we have Windows SharePoint Services, which is the "light" version and comes at no extra cost within the typical Microsoft operating system and Office licences that legal organisations now have. We also have full-blown SharePoint Server 2007. The rapidly developing search capabilities, and the ability to configure and add the necessary metadata to content, are just two reasons why this article suggests SharePoint is increasingly the no-brainer. I am no Microsoft fan, I assure you. I just suggest that here is where we bow to the inevitable.

The point is that your organisation can keep the investment already made in various point solutions, and add to them. This allows you to maintain a best-of-breed policy. For example, your document management system, your client relationship management system, your library management system, your extranet management system, and all the rest, were designed to do a specific job. The organisation that uses SharePoint, to coordinate the core functionality of these applications, may also find that it can then ignore the overlapping pieces that are not core functionality, but can be time-consuming to learn, and configure.

Grey areas

There are two pieces which remain slightly outside this warm and cosy approach. The first is the inevitable matter of the search engine. I am quite glad that I am not peddling expensive search engines today. This is not because the products are not marvellous; they clearly are, and some of them are also good value for money. It's because Microsoft has promised to create, within SharePoint, the MOSS "enterprise search", to rival anything that the current providers can do – but within the existing licensing. The jury is out on whether they will deliver: but for what it is worth, my prediction is that there is too much at stake to miss out on this, and so they will indeed deliver, and, importantly, deliver at the right scale for legal organisations.

If that prediction is right, one point solution that you may not need by 2009 will be specialist search engines. Don't shed too many tears: the core market of the big search engine developers has never been legal: legal

organisations are simply too small to matter that much to them. The largest law firm is tiny compared to the big corporations and government departments who are the main customers for these applications.

The second piece is workflow. Here I am not so sure. SharePoint is increasingly including useful workflow components; but the leading workflow systems can be considered simply as point solutions in their own right, operating within the delivery mechanism (ie the portal functions) of SharePoint. Workflow is so specific to the type of organisation, and the work it does, that this may prove to be the safest way forward. Such fence-sitting is possible for this article, because workflow, as discussed above, is applied to many more aspects of your organisation's work than content management. If you need workflow for content management and you have not already got a specific workflow application, then SharePoint will probably provide it perfectly adequately. If you already have a specific application, used, and almost certainly bought, for other processes, then why not use it also for content management?

The role of the IM professional

Re-useability

The distinctive thing about content in legal organisations, which we have not yet touched on, is the degree of re-useability. This varies from a document template, used by most people several times a week, down to an email acknowledging a meeting: in other words, there is every range between core, multi-use, multi-access material, and the totally ephemeral. But it is quite easy to spot the material which has substantial re-use value, and yet contains

significantly and widely varying actual content. Think precedents, staff manual components, operating procedures, know-how, counsel's opinions – you get the drift.

What distinguishes this re-useable, but distinctive, material is that the organisation makes a big investment in it. It has to be checked against various forms of risk analysis. It has to be kept valid and reviewed. It has to comply. It has to conform. That expense has a corollary: it has to be available, clearly fitted to the exact purpose, to everyone in the organisation who might need it. And the way that is achieved is via metadata.

Making content re-useable

We sometimes hear discussions which almost seem to suggest that "search" and "metadata" are conflicting ways of doing things. This, as readers of LIM will know, is not helpful. "Search" is how you find material in which you have not had to invest much – one-offs, or rarely re-used content – which might be immensely useful in different contexts where a small part of the whole content piece is useful, or in one of the myriad non-recurring situations where information is collated and gathered. "Metadata" is how you find the carefully checked, measured, re-useable material, where context and purpose are clearly constrained.

Metadata – apart from the boring Dublin Core stuff – is the province of the IM professional. It is one of the things you are trained to do. It is a particular skill, and in the legal arena, few lawyers possess it – (why should they? You don't know how to merge banks, after all!) Content management without IM skills in applying metadata will only ever deliver a part of the solution. Learn about SharePoint 2007, if you haven't done so already, so that you can play your proper role once the technical chaps have done their bit.

Biography

Derek Sturdy is a consultant at Tikit. He previously founded and ran Granite & Comfrey, a knowledge management software and taxonomy company and before that he was a director at Sweet & Maxwell with responsibility for the technical implementation of Westlaw UK.