

The Architecture of an Information Revolution

Abstract: The twenty-first century is already proving to be the real dawn of the information revolution. The internet is now officially coming of age and highly functional mobile devices are everywhere. The youth of today are permanently connected and the media world is shifting to a “born digital”, “on line first” mentality. Every part of our society is feeling the effects of this revolution and is reaping the benefits of “real time information everywhere” ... or are they? Adrian Dale, in the second Willi Steiner Memorial Lecture, suggests that the reality is that many businesses are failing to keep pace with information growth and desperately need help – the question is from whom? **Keywords:** Enterprise information management; knowledge management; law librarians; records management

Laying down the challenge

Many twenty-first century business have a serious problem - most of them were not designed for this connected revolution. Their processes and systems were conceived as stand-alone entities not connected within the firm, and certainly not to the wider world. Business models have assumed that the firm is a discrete entity with hard physical and electronic boundaries around it and with staff tied to it for long periods. However, the world has shifted. Most business transactions are now between multiple parties and can be of any size from the 79p of iTunes to multi-billion pound deals. All are mediated by the exchange of complex information in near time and there is little distinction between staff and contractors in any deal.

This brave and exciting new world needs a revolution in the information professions currently working within the firm. The separate churches of IT, library, knowledge management, web services, records management, and ebusiness need to be forged together into a new strategic presence in the firm. And the question for the law librarians of the twenty-first century is ... do you want to lead this revolution, or be led?

The divided organisation – internecine struggles within the profession

The business challenges of information overload are being recognised across the firm at many levels in the

hierarchy. But most firms have not yet evolved the skills to manage the challenges effectively at an individual, team or corporate level. The situation is coming to a head in many different guises and multiple initiatives have been started to solve it. However, the nature of the solutions devised depends very much on the professional background of the individuals involved in the diagnosis. Polarised positions and solutions frequently emerge as different breeds of information professional get involved.

A new holistic approach is needed in which the balance is redressed and the four key communities of our professional domain come together.

These four communities have historically used very different and conflicting terminology to classify business activity. The standards in these areas are now beginning to converge, but many organisations have projects that are still on a divergent path.

Knowledge managers have focused principally on know-how and the organisation of documents/facts/relationships in KM systems.

IT architects have focused on the information systems and supporting technology required to support businesses. Data modelling, object modelling and database design have featured strongly in their work with a focus on very structured information.

Librarians have focused on the unstructured information held in document systems or in physical collections using library style classification terminology

Records managers have focused on the records series produced by business processes and have classified information in file plans supported by additional key words.

As information systems become more sophisticated, there are no longer clear dividing lines between these

requirements and an integrated architecture is required to support them with a language that brings the fields together.

The UK Government has made things more complex by putting in a series of recommendations in each of these areas which, until recently, were conflicting or inconsistent in some respect.

Enter the Enterprise Information Manager...

A new breed of information professional is now needed, someone who can bridge each of these camps and harmonise the language and concepts involved. The Enterprise Information Manager is emerging with a clear understanding of how these concepts all fit together and a good model showing how they can make a powerful and unique contribution.

These managers have a special set of skills. They need to be able to see the organisation differently – viewing it through an information/knowledge “lens” revealing the challenges, problems and opportunities that are presented.

These problems have always been there, but now they are of strategic importance and the Enterprise Information Manager needs to step up to the mark and make her voice heard. The organisation must be made to listen before the problems grow too large to solve.

... Following a well-established organisational development pattern

Some clear development patterns have emerged in the design of organisations during the twentieth and twenty-first centuries. They have evolved through clear phases, each with a significant change characterised by the maturing of a “back room” administrative function into a board level professional discipline. These changes have been driven by increasing sophistication, volume and complexity – all requiring more co-ordinated and skilful management beyond the capability of generalists.

- In the 1950s – Accounts & Finance Departments matured to become the Finance Directorate (run by the Director of Finance or the Chief Finance Officer)
- In the 1970s – The Personnel and Training Departments matured to become the Directorate of Human Resources (run by the Director of HR or the Chief People Officer)
- In the 1980’s – Transport, Procurement, Manufacturing and sometimes Customer Service Departments matured to become part of an Operations Directorate (run by the Director of Operations or the Chief Operating Officer)

Information management is the latest of these, but it suffers from the disadvantage that it is the amalgamation of a range of sometimes collaborating, but more often competing, disciplines:

- ICT
- Records management
- Knowledge management
- Librarianship

The biggest battles have recently been fought between information and knowledge management, when a wide range of professional disciplines entered the fray and moved into the territory traditionally occupied by the librarian.

This has been exacerbated by the maturing of the internet and intranets allowing a new range of communications and media professionals to join the mix. We need a model that accommodates and underpins all of these areas.

We now need the **Information Management Directorate** run by the Chief Information Officer (CIO) or Director of Information. The CIO is a well recognised position in many larger organisations, but is still very technology focused – one of the reasons that the title of Director of Information (which still runs the ICT resources) is to be preferred. The CIO needs to be supported by a new breed of **Enterprise Information Managers**, multi-skilled information professionals who can speak many information dialects.

Quantifying the global problem

The sheer scale of the challenges we face defies comprehension, but let’s try to put it into figures using headlines emerging monthly in the press:

- **“Morgan Stanley fined \$15million for being unable to hand over e-mails demanded by the SEC.”**
 - Computing May 2006
- **“54% of businesses cite compliance needs as being the major driver of storage growth”**
 - “Hot on the Audit Trail” - Computing August 2006
- **“Employees spend 35% of productive time searching for information online.”**
 - Working Council for Chief Information Officers
- **“The Fortune 1000 stands to waste at least \$2.5 billion per year due to an inability to locate and retrieve information.”**
 - IDC, (The High Cost of Not Finding Information)
- **“We are now creating more digital information than we can store, ”**
 - EMC Executive VP Mark Lewis.
- **“This year, for the first time, there won’t be enough storage capacity in the world to hold all the stuff being created”**

- John Gantz IDC

- “161 billion gigabytes of digital content created in 2006, 231 gigabytes for every person in the USA today”

- USA Today – March 2007

This last headline in particular warrants further analysis. In a study in 2003 the University of Berkeley¹ estimated that the corresponding 2002 figure was **5 billion gigabytes**, representing a compound growth rate of 138% per annum since 2002. Or to put it another way – more information was created in 2006 than was created in the whole of the twentieth century!

But what about the situation within the firm?

My own studies show that even the best managed firms have a growth rate of between 50–70% per annum in their internal information volumes. The explosion in unstructured information has been well documented in the IT press and the response to date has been to increase the volume of disk space to cope. However a disk space management crisis is now emerging, with daily back-up times frequently exceeding back-up windows. New IT strategies are being devised to address this and new systems are being installed to bring order to the chaos – with the Electronic Document and Records Management System (EDRMS) taking pride of place in many firms.

However, unless they are well managed, EDRMS systems are actually likely to exacerbate the problem by encouraging the retention of older versions and of project ephemera. EDRMS systems need to be supported by equivalent and integrated systems to handle e-mail, web and intranet content management. Unless this is done, some of the content in the EDRMS will be duplicated elsewhere as e-mail attachments and reformatted web content.

Recent studies by StorageTek² have shown that 80% of information in organisations is replicated material and that less than 10% of information needs to be readily to hand when undertaking retrieval – provided that it is the right 10%. The rest of the material is only necessary for archive reasons, but must be linked back into the main visible collection.

Solving these problems requires all of the information management disciplines to see the problem in the same way and to come together to provide an integrated solution.

Storage is not the goal – retrieval and reuse is

More importantly, finding relevant material in this emerging mess will become akin to the challenge on the world wide web – with one notable difference – the absence of a Google equivalent.

Google has tamed the web by relying on millions of people cross-linking pages manually in a way that drives its accurate document ranking mechanism. The pages that have

been cross-linked the most are usually the most important. Duplicates or older versions are usually ignored in the linking process and drop out of sight on search engines.

Within organisations, there is almost no cross-linking process. Most office documents or e-mails exist as little islands of information divorced from all context. A search engine is as likely to find outdated or inaccurate documents as the most recent version. This is especially true in a world of low information literacy, where most staff have received no training in the art of structured search. Information seeking behaviours and strategies are primitive and often focus on finding **any** relevant information, not necessarily the **right** information.

Editorial skills, librarianship and information science are key parts of the mix

With growing information volumes we need to bring increased order to our information collections, allowing audiences to filter the overload down to manageable levels.

There are four levels at which this can be done:

1. Editorial approach

Information organised by focused “collection” is always the best option if it is well managed and curated. In this arrangement, audiences are presented with documents in context – perhaps the most recent or relevant first – with links to supporting material. This is the driver behind the success of the web as an information tool. A web site which documents the “latest” or “hottest” links is always better than a random hit from a search engine.

2. Classification approach

Yahoo’s early success resulted from armies of students who grouped web information into a series of classes organised in hierarchies. Audiences could find what they needed by navigating through the hierarchies until they found the correct class of information – very much analogous to using a library catalogue with class marks.

3. Categorisation approach

More recently with management information systems and some web based e-commerce systems, we have become used to “filtering”, where we make selections of keywords from multiple (faceted) pick lists and then the system looks for the documents that are classified with all of those concepts combined, often restricted to searching in pre-defined fields. Thesauri can be added to this approach to allow for language variety.

4. Full text search

The Google approach (and other engines) is a free text search where hits are determined by the frequency of occurrence of words in the document – even if they aren’t core to the content. Increasingly sophisticated

algorithms are being devised which improve this approach by using a variety of mathematical techniques to deal with “concepts” and “fuzziness”.

We describe here a sophisticated and managed information landscape, an *architectural* approach that has been carefully designed to meet business goals. Most organisations have evolved an accidental landscape and architecture partitioned across multiple incompatible IT systems. The modern Enterprise Information Manager needs to create the roadmap that will bring order to this chaos.

The five components of the effective information architectures

Most organisations have done a good job in bringing their IT infrastructure under control. However, most ignore (or do not understand) the importance of the *infrastructure*, the *publication and curation processes* and are not aware of the *organisational and governance structures* necessary to sustain the whole architecture.

Whilst these are specialised functions, there is a need to get them understood more widely across the organisation and to build them into the normal performance management processes.

1. IT infrastructure

Many organisations have an excellent IT infrastructure under development, often providing for integrated file store, database engines and a security management infrastructure. This has been driven by the need for systems interoperability and cost control.

2. Infostructure

However, there is often a relatively weak infostructure (the semantic infrastructure) Information can be stored anywhere on the IT systems without regard for classification and metadata and there isn't a high level plan for collection and delivery channel design.

3. Publication and curation processes

To make infostructures come to life – more robust publication and curation processes are needed to assure the quality and accessibility of information stored on the systems and to ensure that the material is delivered to the right audiences in a timely way.

4. Information roles and competencies

This approach requires new organisational roles focusing on enterprise information management to be implemented and new information competencies amongst staff.

5. Governance framework

To make the whole architecture hang together, the organisation needs to adopt stronger information governance policies which curb the current free for

all. Information is a critical organisational asset and needs to be managed with the necessary embedded values and behaviours which reflect the organisational importance of these assets.

The New Enterprise Information Manager's Toolkit

To force (and hopefully lead) the organisation towards a more integrated future, the Enterprise Information Manager needs a clear framework that can be easily explained to business colleagues.

The questions and answers set out below have served the test of time by working successfully at all levels of the firm, allowing a meaningful dialogue to take place on a subject that is often ignored.

1. **For what purpose?** – Information gives our organisation an edge by ... and we measure it by ...
2. **For whom?** – The **Audiences** that we need to engage and satisfy are ... and they need ... in order to do ...
3. **What content specifically?** – The **Collections** we need to assemble and manage proactively are ...and their organisational value is ...
4. **Through what channels?** – Our Audiences access these Collections to achieve their goals through the following channels ...
5. **By what processes?** – We will **Publish** key information to our Audiences, **Curate** our Collections and **Govern** the Enterprise Information Management framework using the following enterprise wide processes ...

Conclusion - the key question

There has never been a more interesting and challenging time to be working in this field. New information challenges are emerging by the day. New information management techniques and tools are appearing even as I write this paper. And all of this change is taking place in the context of an unprecedented and increasing growth in information volumes and complexity.

The business world risks being overwhelmed by this tidal wave of information. And this is just at the point when the law is demanding better record keeping, better governance and better discoverability.

It is the Enterprise Information Manager's job to do something about this. They must lead from the front, marshalling and directing their fellow professionals on many information fronts. They must channel and redirect this tidal wave (we will never turn it back) solving today's information problems and creating tomorrow's business opportunity.

Welcome to the Board Room – your promotion is long over due.

References

- ¹<http://www.sims.berkeley.edu/research/projects/how-much-info-2003/>
²StorageTek 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

Legal Information Management, 7 (2007), pp. 160–164

© The British and Irish Association of Law Librarians Printed in the United Kingdom

doi:10.1017/S1472669607001521

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". Hmm. 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 weaselly 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