

Information Architecture: A Stronger Foundation for Knowledge Management

By Adrian Dale, Creatifica Associates, UK, Visiting Professor of Knowledge Management



A University of California at Berkley Study in 2000 estimated that 250Mb of unique digital content was published in that year for every man woman and child on the planet. Today, the rate of publication is increasing. The world wide web alone is now doubling in size every year and the world's premier search engine Google now crawls and indexes 3.3 billion pages using a cluster of 10,000 indexing servers (statistics from Google at www.google.com). As knowledge management initiatives gather pace, the volume of online information available to employees is similarly increasing exponentially. Disk space usage in many organisations is increasing at a rate of 50-70% per annum (Source: StorageWorld Conference 2003).

In the face of this explosion in information volumes, many organisations are turning to Internet technology to solve their information retrieval problems. Search engine sales are increasing as these organisations look for a way to find the information they need from all the various disparate repositories on their servers. Unfortunately, many of these initiatives are doomed to failure. The popular techniques being used to capture and store corporate knowledge in the form of documents and web pages are simply adding to information chaos by failing to impose any defined structure or context on the materials as they are stored.

Content Replication is a Major Issue

The StorageWorld conference also reported that over 80% of corporate digital data is replicated many times. Not only is this a waste of disk resources, it is creating "information noise" making focussed information and knowledge retrieval much more difficult. Searches often now yield multiple copies of the same document (often alternative versions) each stored in a different repository. For any document, the master copy exists in the file store or document management system. Several intermediate copies are probably stored in the mail boxes of employees as attachments. Yet another version is probably posted to the Intranet and the final version might be posted to the corporate web site or client newsletter as a PDF. Even worse, extracts are now easily taken from any of these various versions to be incorporated into other documents without any indication of their provenance or source. Such a chaotic and unmanaged materials lifecycle would not be tolerated in any other part of an organisation's production system. Raw materials, parts and finished goods are normally carefully managed - tracked efficiently around a factory by an Enterprise Resource Planning system. By contrast, the intellectual output of the firm is often left to fend for itself - stored and managed according to each individual's preferences.

Information Architects - White Knights to the Rescue

The time has come for organisations to pay more attention to the design of information and knowledge storage schemes. It is no longer practical or desirable to have multiple knowledge and information repositories around the organisation, each with a different structure. The relatively blunt instrument of a search engine is not enough to locate the material you need from the multitude of repositories available. Information and knowledge resources are too valuable to leave their storage and capture to chance. These problems are now increasingly being addressed by Information Architects - a relatively new profession dedicated to improving the findability and navigability of enterprise information through more effective repository and interface design. The profession developed initially in the 1980's when information systems became more complex and needed to interact with each other. The field has recently moved into the mainstream of knowledge and information management in response to information navigation issues on the web.

Web Site and Intranet Information Architecture

As web sites increase in size and complexity it becomes more difficult for readers to locate the information they need. Web content today needs to be organised much more logically and yet still creatively enough to grab and sustain the interest of the reader. Site navigation systems need to deliver the shortest possible routes to the most requested content whilst still maintaining a coherent hierarchy. For new Information Architects, Rosenfeld and Morville's (2002) "polar bear book" (so named for its cover) is the definitive text - required reading for all students and very useful for practitioners. This text covers all of the main aspects of information architecture design:

- o Information seeking behaviours (browsing, navigating, searching)
- o Information structuring (content modelling, taxonomies, meta data)
- o Navigation interface design (global, local, contextual, blueprints)
- o Search system design (ranking, clustering, classification)
- o Content life-cycle design (editorial processes, content management)

The excellent implementation methodology described could be applied to any content centric knowledge management project - ensuring that the knowledge in the system is easily navigable and accessible, avoiding information overload.

Good Web/Intranet Design is not enough

Good Internet and Intranet information architectures are certainly important but in most organisations these systems represent only a small proportion of their total online content. Much more information exists in databases, office documents and e-mail - all potentially stores of corporate knowledge. As we have seen, the organisation and management of this material is usually poor - with findability and navigability a major problem. Here a different approach to information architecture is required - one working at a highly level of abstraction and based on the concepts of enterprise information modelling. Elaine and Roger Everden of the 4thResource have focused their work in this area - defining their branch of information architecture more generally:

Information architecture is the foundation for managing information in general as a corporate resource. It describes the theory, principles, guidelines, standards, conventions and dimensions that are necessary to design an effective management framework for information. Its purpose is to design information structures that help people to use information in effective, productive and innovative ways.

Their new book on the subject has just been published (Evernden & Evernden, 2003) and looks destined to become as much of a bible for this branch of the profession as the "polar bear book". The Everdens recognise 8 dimensions of enterprise information architecture, each of which needs to be analysed and understood in any organisation in order for a coherent enterprise architecture to be developed:

1. *Types of information* stored and used across the enterprise
2. *Levels of understanding* required to effectively access and exploit the information and knowledge in the enterprise
3. *Types of representation* defining how the information is structured and stored within the information and knowledge systems
4. *Levels of transition* looking at the lifecycle of the information and knowledge, how it changes over its lifetime.
5. *Types of knowledge* exploring the relationship between information and knowledge in the enterprise and the balance of tacit/implicit/explicit.
6. *Levels of Responsibility* defining content ownership and responsibilities
7. *Types of process* exploring how the various information and knowledge resources will interact with the business processes of the organisation.
8. *Meta levels* overarching the whole architecture and defining holistically how the various information and knowledge repositories interact.

Bridging the Two Schools of Information Architecture

In our experience, the 4thResource approach is a powerful framework for high level enterprise analysis. Equally, the Rosenfeld and Morville methodology is invaluable in the information and knowledge system design phase. However, we have also found a need for an intermediate approach in our work. We undertake architecture layer analysis (Figure 1) beginning with a simple mapping exercise to highlight the current state of an organisation's information and knowledge environment.

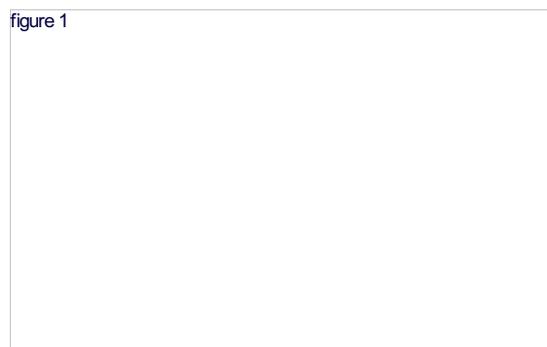


Figure 1. Information Architecture Layers.

We focus on exploring the seven layers shown and the surrounding frameworks:

1. *IT Infrastructure Layer* looks at the numbers and types of servers and storage systems involved together with their naming and the relationship between them
2. *Repository Layer* documents the different content repositories used for structured data, email, documents and web pages.
3. *Content Management Layer* documents the various management tools used to manage the content. This might be a specialised content or document management tool sitting on top of a relational database repository, or it might be a simple arrangement of drive mappings on a server's file system.
4. *Information Model Layer* documents the metadata and controlled vocabulary schema in use across the organisation and any object models used to define more structured data.
5. *Collections Layer* explores the various logical information and knowledge collections that are in use across the organisation to meet the needs of its various business processes. Some information might reasonably exist in multiple collections in different formats but ultimately the architecture should find a way for the content to be stored once in a single location and repository in a neutral format.
6. *Services and Applications Layer* documents or defines the various functions that need to operate on the collections - authoring, depositing, approving, searching, visualising, syndicating etc.
7. *Interface and Navigation Layer* documents how the reader will interact with the various information and knowledge repositories in the architecture.
8. *Publication Process and Curation Framework* defines the lifecycle of the various information and knowledge resources in the architecture.
9. *Organisational and Governance Framework* defines the overall responsibility for the architecture and the local responsibility for each layer.

The maps resulting from this analysis can be powerful tools to drive the development of the information architecture as they highlight the inconsistencies that have evolved in the organisation without it. Figure 2 shows an example of an Information Architecture Map from a UK police force.

In this example, it was immediately clear that the organisation had a repository layer problem with multiple relational database platforms and no information model for unstructured data. There was a powerful data model for structured numerical data but there was a problem with compliance. Some districts of the force had developed non-compliant systems that were then impossible for the aggregation systems in the services layer to deal with. Across the architecture, there had been no investment of time or money in content management resulting in a proliferation of unconnected information and knowledge resources. This was compounded by the absence of a coherent publication process and curation framework and by the weakness of the overall governance framework. With a diagram of this nature and a few choice words it doesn't take long to convince a client that they have a serious problem.

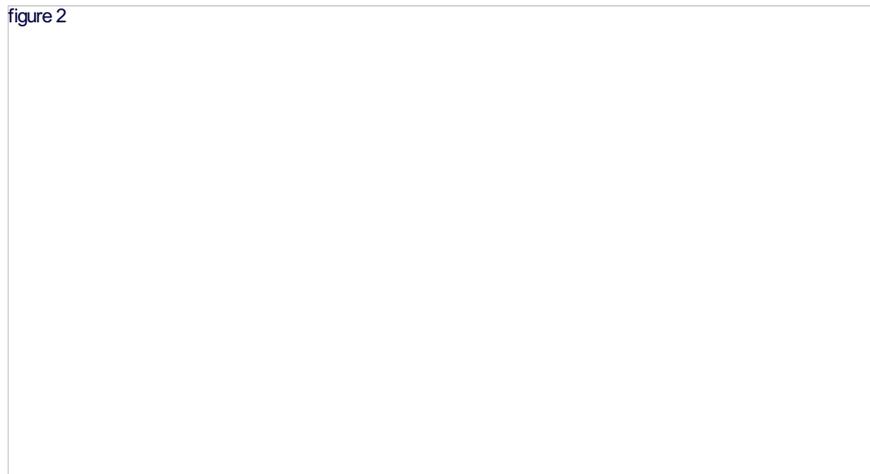


Figure 2. Example of information architecture map.

Concluding Remarks

The re-emerging field of information architecture has much to offer the knowledge management movement. It provides the thinking, the methodologies and the tools to make the knowledge management dream scalable - enabling it to deal with the countless terabytes of online information that will be flooding the world wide web and corporate information systems over the coming years. Without information architecture, information overload could well bring the whole knowledge management crusade to a halt - and we'll all be worse off.

References

- University of California at Berkeley (2000), *How Much Information?*, available at <http://www.sims.berkeley.edu/research/projects/how-much-info/>
- Rosenfeld, L. & Morville, P. (2002) *Information Architecture for the World Wide Web*, O'Reilly Publishers.
- Evernden, E. & Evernden, R. (2003), *Information First: Integrating Knowledge and Information Architecture for Business Advantage*, Butterworth-Heinemann.