Knowledge Architecture:
A Vision for the 21st Century

R. J. Bater
KnowPlexity Ltd.
43 Ashley Down Road
Horfield
Bristol
United Kingdom
bbater@knowplexity.com

Summary

The realization that information and knowledge are key resources in the 21st century knowledge economy, and that they need to be managed, came late to many organizations. Few were well-equipped to face the challenge. For many, the result has been a 'BandAid' solution, where inadequate commercial off-the-shelf software has been applied to an ill-defined set of requirements at no small cost.

Two key problems facing organizations in this situation are identified. Firstly, the great majority of applications offered as 'knowledge management solutions' are actually information management solutions and fail to deal not only with the problems of managing email, instant messaging and Web 2.0, but also to engage with the essentially human behavioural and cultural issues. Secondly, understanding seems to be lacking among both vendors and practitioners regarding the complexities which can arise when trying to 'manage' this human feature called 'knowledge'. Leading thinkers in the field are drawing upon the unlikely combination of complexity theory and anthropology to forge a new paradigm of knowledge and information creation and transfer. Yet, this work is not widely known.

It is proposed that the relatively new discipline of Knowledge Architecture promises a heterogeneous, holistic framework within which information management can evolve to meet the challenges presented, by recognizing that organizations are complex adaptive systems and by the application of tools and techniques such as sensemaking and social network analysis.

Key words: knowledge sharing, information management, knowledge management, knowledge architecture

Introduction
IM Awakening
A report by IDC in March 2008 (IDC/EMC, 2008) estimated that the digital universe in 2007 comprised 281 exabytes (1 exabyte = 10^{18} bytes), and that by 2011, the digital universe will be 10 times the size it was in 2006. Most people look at figures like this and see a threat. It portends challenges, certainly. But information professionals should also see opportunities.

If information production is growing so fast, then information must be important. So, information matters! And that happens to be the title of a report published by the UK government last November (Knowledge Council, 2008). In the Information Matters report, Sir Gus O'Donnell, the Head of the UK Civil Service and Cabinet Secretary writes in his introduction:

"Good information management needs to be partnered with good knowledge management. If it isn't, the value of information as an asset is undermined, and cost-effective, efficient service delivery is compromised."

Yet, only the year before, surveys revealed that:
- on average, staff spend 9.6 hours per week searching for information (IDC/EMC, 2008)
- at any given time, 3% – 5% of files can't be found. Recreating them costs on average €100 per document (Information Week Survey, 2008)

There is obviously a mismatch here between the aspirations of Information Matters and the reality of the surveys. If information professionals want to rectify that mismatch, that presents them with a considerable challenge. It's a challenge they welcome, of course, but to meet it on its own terms it will take time, resources, and a better understanding of how knowledge and information *together* generate value.

BandAid Response
We long ago responded to the problems of managing money by devising complex systems for managing it. Information has not fared so well, and the response has proved to be piecemeal and blinkered, with each function attending to its own needs. When computerized information systems began arriving in large numbers in the 1960s, every information problem was tackled independently. In the absence of any overall scheme, the result was a collection of ad hoc, single-purpose tools, each serving their own organizational silo:
- a financial accounting system
- a customer database
- an employee database
- a payroll application
- management information systems

Today, we call that 'Information Scatter', that is, an organization's information is scattered among multiple, incompatible, non-communicating information

systems and data stores. A whole new generation of middleware has evolved in recent years to address this problem, like IBM's WebSphere application integration software, or Schemalogic's suite of applications to cross-map and manage metadata across multiple, siloed information systems. Such responses can be viewed as an attempt to derive order *a posteriori* out of a chaos that need not have existed in the first place. The result is complete bewilderment when users need to find specific information.

Holistic Health

The problem is, knowledge needs information and information needs knowledge, a truism that seems to have been largely ignored by the recent histories of both information management and knowledge management. Without information to feed it, we can have no knowledge. And without knowledge, information is useless.

The 17th century English philosopher and scientist Francis Bacon is often quoted as saying "Knowledge is power." But this is almost a complete inversion of what he actually said. The quotation in full is:

"But mere knowledge is not power; it is only possibility. Action is power; and its highest manifestation is when it is directed by knowledge." (Bacon, 1597)

So, knowledge is the power to take effective action.

Problems in Knowledge Architecture

Information + Knowledge: A New Kind of Problem

A knowledge architecture is, if you like, an expression of the 'personality' of an organization - what it exists to do and for whom, how it goes about doing what it does, who does what, and what information and knowledge are used - and produced - along the way. Building a knowledge architecture which will work well for an organization, is therefore a bit like being a psychoanalyst. You have to delve deeply into motivations, the rationale behind patterns of behaviour, the influence of legacies from the past, whilst being prepared to come across all sorts of disorders and complexes lurking beneath the surface, just waiting to be discovered.

What's really needed is a framework where the jig-saw of business processes, information systems, information resources, and the knowledge in people's heads, can be put together to make a meaningful picture. That's the challenge of Knowledge Architecture. There are problems with information, and there are problems with knowledge.

Information in the Wild

The problem of information scatter has already been mentioned. In reality, the information fragmentation problem is far worse. Much of the information we need to capture is wild. It doesn't reside in structured environments like data-

bases and libraries. It grows where and when it's needed – and often where it's not.

A recent working paper by UK National Archives CEO Natalie Ceeney estimated that 80% of an organization's information is of this type (Ceeney, 2009). Wild information resides in documents, emails, instant messages, blogs, wikis, del.icio.us, Technorati, Twitter – the list could be extended.

While some documents remain native to paper, most are now born-digital: Microsoft Office, Open Office, HTML, XML. We have EDRM systems for managing Office documents, but HTML and XML present a whole new set of problems. Email steadfastly refuses to limit its growth and to be manged, while instant messaging presents further problems of its own. Both almost certainly contain vital information. With Web 2.0 applications – blogs, wikis, URL collections like del.icio.us and Technorati – we must ask how we might identify what's important, how we might know when to capture something which is continually changing, and how to organize it? Some Web 2.0 commentators like Clay Shirky (Shirky, 2006) say that conventional information management techniques like classification and taxonomy are irrelevant anyway for Web-based resources.

Taming Wild Information

Even if we can capture what needs to be captured, we cannot assume that information from different sources can be easily aggregated without further cleansing, processing and mapping. Lou Rosenfeld, a founder of the North American Information Architecture movement, illustrates these problems as shown in Table 1 (after Rosenfeld, 2003).

Table 1: Metadata interoperability problems

Source A	Source B	Source C
Metadata are not interoperable		
Create Date: 071005	**Date**: July 10, 2005	**Created date**: 100705
Author: William Jones	**Compiler**: Bill Jones	**Creator**: Jones, W.
Subject: guidelines	**Topic**: policies	**Descriptor**: procedures
Structural interoperability achieved through applying standards		
Date: 071005	**Date**: July 10, 2005	**Date**: 100705
Creator: William Jones	**Creator:** Bill Jones	**Creator:** Jones, W.
Subject: guidelines	**Subject:** policies	**Subject:** procedures
Semantic interoperability achieved via controlled vocabulary		
Date: 2005-07-10	**Date**: 2005-07-10	**Date**: 2005-07-10
Creator: Jones, W.	**Creator:** Jones, W.	**Creator:** Jones, W.
Subject: policies	**Subject:** policies	**Subject:** policies

Even internationally accepted classifications can pose problems. An EU institution had to deal with other economic institutions all over the globe. They used ISO 3166 (ISO 3166, 2006), which defines the international standard codes for representing countries, but it only supported some of their needs. For the rest, they had to improvise.

ISO 3166 is a geographical vocabulary, and the institution has to work with supranational, geoeconomic and other groupings as much as the geographical. For example:

Table 2: Geographical vocabulary problems

Supranational Groupings	Geoeconomic Groupings
Near East	COMECON
Middle East	EFTA
Far East	OPEC
Europe	Eurozone
Development Groupings	Legacy Geopolitical Groupings
ACP (Africa, Caribbean & Pacific)	FYROM – Former Yugoslav Republic of Macedonia
ALA (Asia & Latin America)	

So, for instance, you won't find supranational groupings in ISO 3166, like Near East or even Europe. And you won't find geoeconomic groupings like COMECON, EFTA or Eurozone either. Nor will you find common groupings used in development economics like those loose alliances of developing countries such as ACP and ALA. And of course, political groupings break up and reform. So the institution struggled to decide whether it should index documents under 'Macedonia' or 'FYROM'.

If we think these are substantial information management problems - and they are – how much more of a problem is it to try to manage something which is intangible, only partially controllable, highly personal and ever-changing? Well, the answer is 'we can't'. But we *can* encourage its growth and its flow, and. Web 2.0 applications offer new ways of doing that. They don't, however, lend themselves well to 'being managed'.

Yet, blogs & wikis only carry knowledge that gets written down. Alongside these 'visible channels' there are other, informal channels - like the telephone, conversations at the coffee machine or in the corridor - which carry a vast daily flow of unwritten knowledge which remains entirely untouched. We can't capture that knowledge directly and shouldn't try. But we can find out who talks to whom, and for what reason. In many cases, that will be enough.

Accommodating Diverse Points of View

Just as Architecture seeks to create spaces for defined purposes, so Knowledge Architecture must seek to create knowledge spaces for defined purposes. Knowledge and information are, of course, used for many different purposes in an organization. This paper proposes that these purposes can be clustered into five overlapping domains, each of which makes different demands of an architecture in terms of the attributes of information and knowledge they view as important.

Figure 1: Domains of use of knowledge and information
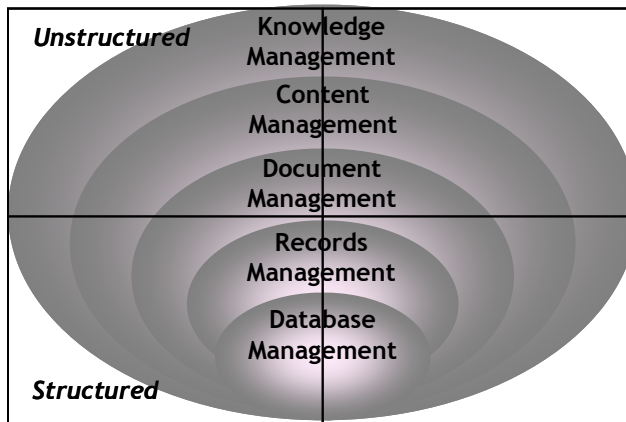
In Figure. 1, from bottom to top, the degree of structuring decreases The purpose of Database Management is to capture, store and make data retrievable. Database Management Systems are, of course, highly structured. The attributes which matter most are things like data type, cardinality and whether a column is a primary or foreign key.

Records Management has as its core purpose, the capture, secure storage and provision of controlled access to a subset of organizational information deemed to be records of the organization's activities. Much of its content will be derived from clusters of documents. The important attributes here are the organizational activity which produced the record, the applicable lifecycle (retention - preservation/destruction) as may be dictated by regulatory authorities, and appropriate access control. Structure is therefore essential.

Document Management, like Records Management, deals with information linked to defined organizational processes, but applies very different lifecycle parameters and access controls. Governance is driven more by internal needs for due diligence than by external regulators. Document Management focuses on the collaborative generation of discrete packages of unstructured information (documents) in the context of a defined organizational process. Controlled distribution, version control, approval workflows and sign-off procedures determine the important attributes and degree of structure.

28

Although not widely recognized as such, the main purpose of Content Management is publication, either to an internal or an external audience. Content Management people therefore need to maintain awareness of all organizational information which can and should be disclosed to defined audiences. Content Management is driven either by the requirements of applicable regulations, by the need to keep staff aware, or by marketing, promotional or public relations considerations.

Finally, the Knowledge Management domain - the least structured - is concerned with creating the conditions in which people can use and share the knowledge they have, combine it with knowledge shared by others, and fuel creativity by drawing on the information preserved in the organization's repositories.

Tools and techniques for the lower four domains are well developed but have now to be applied organization-wide, with all of the classification and metadata interoperability problems that implies.

But the tools and techniques for achieving the aims of knowledge management so far offered barely break the mould established by information management.

A Vision for Knowledge Architecture

We need to look elsewhere for guidance on how Knowledge Architecture can accommodate knowledge alongside information and facilitate its 'management'; how it can be inclusive of all points-of-view on the organizational knowledge base and the diverse uses to which it is put. A number of innovative analytical frameworks - have emerged in recent years which may offer a solution to the impasse. Three, of particular significance are summarized here.

The Cynefin Framework

When Dave Snowden was Director of the IBM Institute for Knowledge Management a few years ago, his research into the relevance of Complexity Theory to Knowledge Management produced what he called the Cynefin Framework (Cynefin Framework, 24 August 2009). Snowden has since developed this framework into a sophisticated methodology for what is known in management epistemology as 'sensemaking'.

The Cynefin Framework defines five domains in which we interact with experience, use information, and apply our knowledge:

- Simple
- Complicated
- Complex
- Chaotic
- Disorder

Cause-and-effect are collective truisms in the Simple domain, supporting logical decisions and best practice, but in the Complicated domain, cause-and-effect are not obvious and good practice is only revealed with effort or expertise. The

Complex domain affords no sense of predictable cause-and-effect, but it may be recognized in retrospect as the source of 'emergent practice'. The two remaining domains exhibit no discernible cause-and-effect relationships. Although 'novel practice' may reveal itself in the Chaotic, in Disorder, decisions will be made largely on intuition.

The Cynefin Framework helps us to understand how we need to define 'knowledge' according to the applicable domain, and to apply different approaches to 'managing' it. Sometimes, 'managing' means making no intervention at all. Instead, we must simply provide supporting systems, and 'Let it Be'.

The I-Space

In parallel with Snowden's work, Prof. Max Boisot, Professor of Strategic Management at the Birmingham Business School in the UK, proposed a model of information diffusion he calls the 'I-Space'. It is regarded as a seminal work in the realm of corporate anthropology (Boisot et al., 2007).

In the I-Space, the efficiency and effectiveness of transfer of information or knowledge is characterized along three key vectors, Concrete-Abstract, Uncodified-Codified, and Undiffused-Diffused. Concrete knowledge is purely experiential, unrefined, and contaminated with 'noise', but may be 'cleansed' and condensed to the abstract. Uncodified knowledge has little structure and requires effort to extract useful knowledge; codification makes it far more accessible.

Abstract, codified knowledge is most easily diffused and converted into organizational value, but abstraction and codification dilute semantics and may prove self-defeating in some cases. Differing degrees of uncertainty and risk therefore accompany certain types of knowledge, and this can have profound effects not only on how people use the different types, but also on the type of people motivated to use them.

The I-Space manifests itself in numerous ways, one which Boisot himself has studied being management culture. Boisot proposes four domains - POSSIBLE, PLAUSIBLE, PROBABLE and ACTUAL - which managers can inhabit in different proportions. As probability of outcome increases when one traverses from the POSSIBLE to the ACTUAL, so risk decreases. Different management styles crystallize, Boisot says, into two distinct types, the managerial and the entrepreneurial, each with its own mode of applying knowledge. It is difficult not to see the parallel between Snowden's four Cynefin domains and Boisot's.

The Snowden and Boisot frameworks can be seen as two similarly-structured sensemaking spaces addressing different, although related sets of referents. At their point of overlap, they share three interacting concepts: organizations, knowledge and people.If the Snowden and Boisot frameworks may be used to understand the processes by which knowledge acquires significance and the processes by which significant knowledge is made accessible, then an understanding of how such knowledge is acquired and exchanged would add a further dimension to any knowledge architecture.

30

The developing practice of Social Network Analysis (SNA) offers much in this respect. SNA uses a variety of methods to identify the nodes (people) active in a knowledge-sharing network, and to characterize their contributions in terms of various roles and types of 'ties'. For instance, some may act as gateways for knowledge flow between one cluster of nodes and another cluster; others may act as hubs, to which much of a cluster's knowledge gravitates.

Conclusions

Knowledge Architecture then, has not only to integrate a multiplicity of forms and formats of information - documents, emails, instant messages, blogs, wikis - encodings like HTML, XML and numerous proprietary encodings from Microsoft, Open Office, and tens of others, but also has to cope with data representation standards, semi-structured and quasi-managed collections (blogs and wikis) and the special needs of specific points-of-view onto the organizational knowledge resource. Knowledge Architecture also has to engage with epistemology and account for what knowledge is, how we acquire and share it and why it is important.

Recognizing that organizations are complex adaptive systems, employing sensemaking techniques to understand how useful activity is sustained by knowledge and information, and identifying the social networks through which much of it flows, will become essential components of the knowledge architect's toolkit alongside the conventional tools of information management.

References

Bacon 1597. Bacon, Sir Francis. Meditationes Sacræ. De Hæresibus. (1597)

Boisot et al., 2007. Boisot, Max H., MacMillan, Ian and Seok Han, Keyong. Explorations in Information Space: Knowledge, Agents, and Organizations, 2007. Oxford University Press

Ceeney, 2009. Ceeney, Natalie. Integrating Information Management into Business Processes: Project outcomes. The National Archives, undated. Available at: http://www.nationalarchives.gov.uk/electronicrecords/records-tools.htm

Cynefin Framework, 24 August 2009. Available at: http://en.wikipedia.org/wiki/Cynefin

IDC/EMC, 2008. The Diverse and Exploding Digital Universe: An Updated Forecast of Worldwide of Information Growth Through 2011. IDC/EMC, March 2008.

Information Week Survey, 2008.

Information Week, March 3, 2008.

ISO 3166, 2006. ISO 3166-1:2006. Codes for the representation of names of countries and their subdivisions - Part 1: Country codes. International Standards Organisation.

Knowledge Council, 2008. The Knowledge Council. Information matters: building government's capability in managing knowledge and information. Government Knowledge & Information Management Network, November 2008.

Rosenfeld, 2003. Available at: http://www.louisrosenfeld.com/home/bloug_archive/images/031019.pdf

Shirky, 2006. Ontology is Overrated: Categories, Links, and Tags. Available at: http://www.shirky.com/writings/ontology_overrated.html