
**A New Sales Recipe for the Enterprise:
Governance, Semantics, and a Dash of Open Source Sauce**

In a meeting with some Central European entrepreneurs, I jotted down a statement one of the people made:

The large enterprise software companies are making up problems as they go along. Even worse, these companies are ungovernable. Their internal information is unmanageable.

I thought about the observation on the taxi ride back to my hotel. At first I thought that it was grousing by a European business executive. After some additional thought, I decided that in certain technology initiatives, the comment was quite accurate.

The consumer and trade coverage of the governance “issue” at Hewlett Packard underscores a blended problem in management, of an organization and its information. Bloomberg’s “HP’s Woes Accelerate CEO Succession Crisis” <http://www.bloomberg.com/news/2011-09-22/hewlett-packard-shares-reeling-47-accelerates-ceo-succession-crisis-tech.html> is an example of swinging for the fences or trying to replace a jet engine when the air craft is in flight. If a Board of Directors can veer out of control on the slipperiness of governance, it should be no surprise that enterprise software systems can be difficult to control. With revolving doors for senior management, does an information policy have much of a chance? Management, not software, determines whether an organization has its act together.

Leo Apotheker’s vision for Hewlett Packard was to convert it to an IBM or an SAP. There is a sharp contrast between the hard specifics of consumer products and mobile phones and what IBM sells is evident in IBM’s description of its content and document management services, adding jazzy words and fuzzy phrases. The updated IBM Web site now includes the Watson search system, built in part on open source technology.

As you may know, Watson is the focus of one of the most interesting public relations campaigns in the last two years. First, the search system “won” the television game show *Jeopardy!* by answering questions only a smart human would know. Then Watson was positioned to revolutionize health care. Now the system is listed as a core IBM enterprise technology. Is marketing and puffery taking precedence over basic information blocking and tackling?

You judge. Here’s what IBM says:

IBM Enterprise Content Management (ECM) helps companies make better decisions faster. Enterprise content management waves of innovation. Advanced case management. Software powered by collaboration, content, process, analytics, and business rules to provide an advanced 360-degree case view. Document imaging and capture. Software for automated capture, production imaging, report management, and standardization and consolidation. Social content management. An end to end platform to manage office documents, web and social content. Content analytics solutions. Software to help you derive new business insight rapidly. Find, assess, analyze, and explore all of your enterprise content. Information lifecycle governance. Integrated software for smart archiving, e-discovery, records management, and disposal and governance management. Source: <http://www-01.ibm.com/software/data/content-management/>

What is fascinating is that “governance” seems to be a slam dunk—no big deal. License IBM software and content is ship shape and on deck. Is governance software? Google told me there were about 125 million Web pages about the term. Wikipedia’s lengthy entry for “governance” <http://en.wikipedia.org/wiki/Governance>. wanders. What I find interesting is that the term “governance” embraces everything and nothing; for example, “decisions that define expectations, grant power, or verify performance.”

How would you answer this question if asked it in a meeting, “Do existing enterprise systems manage digital information effectively?” Content management, text analytics, and other “value adding” information processes reveal the management flaws in essential processes. Buzzwords and puffery or an accurate depiction of reality?

My view is that “governance” is similar to a flexible, polyethylene sheet that can be wrapped around an organization’s digital content output. At first glance, setting up rules, guidelines, procedures, and workflows seems air tight. Using software that implements content controls looks like a prudent decision. However, like a chicken leg swathed in plastic wrap, the inevitable is just delayed. Chicken and good intentions spoil. CMS is the polyethylene that allows an observer to watch the information process. CMS does not remediate an information process.

Governance, in my opinion, when applied to digital information is a subset of corporate management work. Specifically the key checkpoints are ensuring that content is produced for a Web site or other “output” when it is needed. The information is vetted, usually through a process of having approvals prior to the publication of the information. Subordinate functions embrace versions of the document, transformation functions so a Web article can be “repurposed” for a PowerPoint presentation, and, of course, finding a particular document.

In short, “governance” means putting in place software, systems and procedures to prevent the digital system from losing control of information. A moment’s reflection about the popularity of Facebook, the immense reach of a Twitter “tweet”, and the marketing opportunities presented by social media make governance increasingly important.

At a meeting in Europe in early September 2011, I learned about TEMIS and its tie up with Alfresco. My information was lunch time chatter, but I spotted a news release on September 8, 2011, which provided confirmation of the table talk and revealed several interesting items.

TEMIS is one of the many European companies pushing into semantic content enrichment. I reported on developments at Spotter in one of my 2011 Information Today columns, and I have documented the semantic technology of Exalead, now owned by Dassault Systèmes in my *The New Landscape of Enterprise Search*, published in June 2011. TEMIS is similar in some ways to the business intelligence centric Spotter and to the multi-layered Exalead CloudView system. But there are some important differences, and I think these make the Alfresco deal even more interesting.

TEMIS, founded by an IBM researcher, Charles Huot, seemed promising and of particular usefulness to those wanting to extract insights from medical and pharmaceutical-related content. That was in 2000, and now 11 years later, TEMIS has expanded its core content enrichment solution which automatically associates value-added metadata to content. The product is called Luxid, which is not to be confused with the open source search system from Lucid Imagination. Luxid's technologies include semantics, sophisticated numerical recipes, and some vertical solutions.

The firm's clients include such organizations as the American Association for the Advancement of Science, Agence France-Presse, BASF, Bayer Schering Pharma, BNA (Bureau of National Affairs), Editions Lefebvre-Sarrut, Elsevier, Merck Serono, Nature Publishing Group, and Thomson Reuters, among others.

Alfresco, as you may know, is an "open platform" for social content management. With nearly 2,000 customers in 55 countries, Alfresco asserts that it is the world's most trusted open platform for highly scalable, enterprise-class content management. The company has both "open" goodness plus a line up of more than 250 partners. In addition, the firm's customers include Home Depot, Michelin and the New York Philharmonic.

Both companies offer high-value solutions. Both companies have an impressive line up of customers. Both companies have a partner network. Both companies embrace the notion "open source" for both technical and marketing reasons. But the interesting facet of the tie up is that a solution to some thorny information management challenges will be available to prospects and customers.

According to the news release issued on September 8, 2011:

The benefits of this Luxid®-Alfresco integration can be felt enterprise-wide, and most notably by making end-user access to relevant content both faster and more effective, bringing productivity and insight to all decision-making and innovation processes, alleviating the need for time-consuming, manual metadata contribution, helping Content Management and Information Management teams optimize their content management, archival and distribution decisions and deal with the growing mass of available content. Source: <http://www.prnewswire.com/news-releases/temis-and-alfresco-unveil-joint-integration-129445908.html>

What strikes me is that TEMIS is describing features and services that echo IBM's language about content management.

The track record for CMS vendors has paralleled the trajectory of enterprise search, semantic tagging, and natural language processing; that is, these specialized functions are now at risk of commoditization. Content management is in the same pickle. A number of CMS vendors have been purchased by companies looking for new customers and sources of revenue.

The logic of the deals perches upon an assumption that buying customers of one software system will set the stage for selling that company an unrelated product. EMC purchased Documentum and recently embarked on a quest to find a search solution that would make content more findable when housed on an EMC storage system and queried with an open source search solution. One of the grandfathers of enterprise information systems is OpenText. OpenText owns BRS Search, Fulcrum, an SGML search system, Nstein, and LiveLink. OpenText acquired RedDot and that system arrives at the customer location equipped with a version of Autonomy's search system. Is this governance in action?

As the high profile CMS vendors have been absorbed into other companies, CMS has spawned a number of open source solutions. Among these open source vendors are Drupal, Joomla, and a mind boggling number of variants. You can get a league table from Wikipedia. Like other Wikipedia listings, I usually double check some of the specifics, but the listing at http://en.wikipedia.org/wiki/List_of_content_management_systems provides a very useful aerial view of a very complicated range of product offerings. Alfresco, like the Hippo CMS which I found quite interesting, is listed in the open source category. Even Microsoft, owner of SharePoint and the Fast Search & Transfer technology, officially endorses Wordpress, a quasi-open source content management system which can support blogs as well as "traditional" Web pages.

Three questions come to mind:

First, with both proprietary solutions like the ubiquitous Microsoft SharePoint and the quite popular Wordpress CMS, is there sufficient market appetite for the more than 50 different systems listed, recycled, repackaged, and shaped by a menagerie of vendors? My view is that attrition is likely to accelerate.

Second, with the Web going on 20 years in age, why is content still such a problem that it requires the large number of vendors whose systems seem, at least on the surface, to perform

similar tasks? I see it as content-centric vendors grasping at any reasonable straw in order to generate revenue, ignoring for the most part the cause of the content crisis in organizations.

Third, is the crisis in governance, which I define as editorial policy, caused by the introduction of a CMS? I hold the view that a content management system exposes problems which were previously invisible.

What can industry partnerships deliver? For an organization with a content problem, partnering can provide a more automated approach to an on point solution. However, some vendors may find themselves offering a modern version of this observation made in *Popular Mechanics* in 1949:

"Computers in the future may weigh no more than 1.5 tons."

In today's business environment, some information issues may indeed be ungovernable and unmanageable. HP cannot become IBM. Partnerships cannot change CMS into governance solutions.

Stephen E. Arnold is a consultant. His most recent book is *The New Landscape of Search*, published by Pandia in June 2011. His Beyond Search Web log is at <http://www.arnoldit.com/wordpress>.