

Abstract

This article discusses the reasons why an open approach to information is required and the motivations behind the decisions by the authors of Method for an Integrated Knowledge Environment 2.0 (MIKE2.0) to hand their "creation" out to the broader professional community.

During the 1990's the volume of raw data held by enterprises has grown tremendously. All of that data had to be put to some use, and it has been both internally and externally. As a result, non-ledger data has taken on greater and greater importance in the management, oversight and assessment of companies. Unfortunately, the use of agreed processes and standards for the aggregation, measurement, quality and interpretation of the data has not moved at the same rate with every enterprise free use their own approaches. In some cases this results in innocent ambiguity while in other cases organizations have taken the opportunity to deliberately mislead their stakeholders.

The complexity of data is not generally well understood. Most often, it is assumed to be a set of static datasets which can be related to each other in an unambiguous way. The reality is that

data is constantly changing across the enterprise 24 hours a day. With financial reporting, this constant change is generally well managed with ledger aggregation, group reporting and, most importantly, period-end closing. By agreeing to specific cut-offs a point of reconciliation stabilizes all of this ongoing change. Although it is taken for granted, the process followed to stabilise the data are non-trivial.

If non-ledger data is to be trusted to the same extent as financial data, then its complexity needs to be equally well managed in ways which are consistent across the industry. No one consulting firm and no one financial institution can find the "right" answer unless the approach is much more widely adopted. For this reason BearingPoint not only invested heavily in developing approaches to managing and measuring complex data, they have donated it to the wider profession using a Creative Commons licensing model.

MIKE2.0 is that initiative and is larger than any one group of professionals. It is managed by a mix of industry professionals across end-user and consulting firms. It is designed as a multi-lingual collaboration that can link external reporting minimum standards with multiple internal data consolidation processes using a variety of technologies. MIKE2.0 is one of the initiatives that Information Management professionals looking to shape their industry can embrace, influence and extend.

MIKE2.0 – The case for an open approach

Up until the end of the 1980's the predominant business model for large complex enterprise was distributed. This was a practical reality of managing a large number of people and servicing sophisticated dispersed customers. The only things that were typically consolidated was the group financial results.

Not surprisingly, decisions at an enterprise or group level focused on those financial metrics and external stakeholder reporting (market, regulatory and political) were heavily biased towards those same measures. The control, quality and standards surrounding the consolidation of financial information sat (and generally continues to sit) with a finance function who adhere to accounting standards appropriate to their jurisdiction.

Between 1990 and 1992 the price of storage fell sharply with new disk technologies and the price per megabyte approaching US\$1¹. Suddenly it became practical to hold more than the minimal data associated with each individual transaction. While the opportunities seemed endless, problems started to appear even by the end of the decade.

Drowning

“Water, water, every where, Nor any drop to drink.”ⁱⁱ. With all of this data many people assumed that by this decade information would be ubiquitous. For instance, Michael Saylor (MicroStrategy CEO) talked during the latter part of the last decade of a “Query Tone” which was going to be simply a matter of tapping into the network of data.

The trouble is, data may exist on a network, but it has not traditionally been networked. Relational theory, first proposed by Ted Coddⁱⁱⁱ identified that individual lists of data have relationships which can be described in an “Entity Relationship” diagram. Going further, Bill Inmon and Ralph Kimball authored seminal books “Building The Data Warehouse” and “The Data Warehouse Toolkit” respectively which argued that, regardless of how the data was generated, it needed to be consumed in an integrated way.

Group consolidation of data means moving from individual black box systems which are in silos to integrated, enterprise and open platforms. Such a move requires a focus on simplifying the data to enable humans to be able to interpret the results. Even if developers are engaged build reports, individual user groups need to directly control and sponsor their own queries. This is really an application of Small Worlds theory^{iv} which shows that every aspect of the data “network” should be unambiguously linked through a small number of steps.

The need for standards

With all of this rich and exciting raw data available to business users, it’s no surprise that they want to use it. In the absence of the managed, relational enterprise view described in the previous section, qualitative decisions are made every day about how different datasets will be integrated and interpreted.

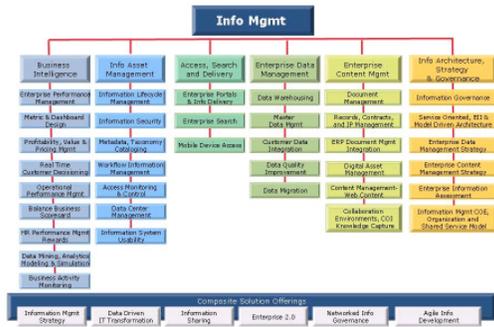
This no different to the qualitative and interpretive decisions made by accountants when integrated different ledgers. However, in the financial case, there are accounting standards to fall back on. Further, there is a whole auditing infrastructure in place to give users of the data confidence that the process has rigor.

For non-ledger data, no such infrastructure exists and there is almost a complete absence of standards across the industry. This means that a report can be presented claiming, for instance, 1 million customers using a range of products without any explanation of whether that is 1 million active customers, whether a customer counts as a family or as an individual, whether they are business or personal etc.. While this may not be important for the majority of situations, when a mistake is made there is no reliable or rigorous method to detect it with the first sign of trouble often coming from outside the organization.

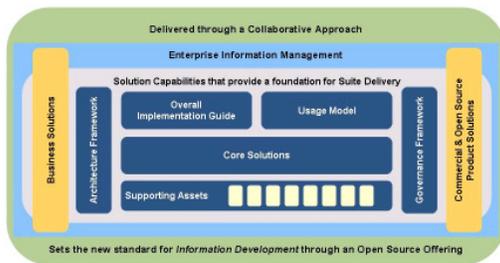
- “BigPond seeing double” (“The Australian”, June 28, 2006): “TELSTRA has been double-counting at least 250,000 subscribers to its BigPond internet service”.
- “RedEnvelope Cuts Outlook, Shares Fall, CFO Eric Wong Resigns Amid Budget Errors” (AP, New York, 2005): RedEnvelope spokeswoman Jordan Goldstein said the budgeting error was simply due to a number wrongly recorded in one cell of a spreadsheet that then threw off the cost forecast.

MIKE2.0

Like many international consulting firms, BearingPoint has developed a set of intellectual property describing how Information Management problems should be tackled. It has released some of this content to the open source community through the MIKE2.0 Methodology. MIKE2.0 covers the complete scope of Enterprise Information Management: Business Intelligence, Information Asset Management, Access, Search and Information Delivery, Enterprise Data Management, Enterprise Content Management and Information Strategy, Architecture and Governance.



MIKE2.0 provides a complete delivery framework Enterprise Information Management projects. It a set of Solution Capabilities that provide a common competency across all provides through an Overall Implementation Guide, Usage Models and shared Supporting Assets.



MIKE2.0 Solutions bring these capabilities together to address the most frequently faced issues in Information Management. Whereas much of the content is focused on “logical” best practices and competencies for Information Management, product-specific solutions are also available on the site.

the enterprise. The approach followed to build this repository is referred to as a [Mashup](#).

As good as BearingPoint believes its approach is, having proprietary approaches to Information Management provided by each of the major consulting firms is analogous to first London fire brigades. Following the “Great Fire” of 1666, insurance companies individually funded private fire services to protect just their own clients’ buildings. Today, if a company has an Information Management “crisis” (like one of the embarrassing examples described earlier) the first step is to work out which proprietary Information Management governance framework is in place.

Standards are not enough

The complexity of data is not generally well understood. Most often, it is assumed to be a set of static datasets which can be related to each other in an unambiguous way. The reality is that data is constantly changing across the enterprise 24 hours a day. With financial reporting, this constant change is generally well managed with ledger aggregation, group reporting and, most importantly, period-end closing. By agreeing to specific cut-offs a point of reconciliation stabilizes all of this ongoing change. Although it is taken for granted, the process followed to stabilise the data are non-trivial.

While standards are important for accounting, it is also the methods, techniques and logical architectures that have evolved and are agreed amongst the profession after more than two centuries of evolution.

For the Information Management profession to even begin to catch-up, standards are needed for key data and role definitions but more than that the terminology and practices surrounding the handling of data in general needs to be agreed.

Financial systems were developed for a profession who understood the inputs, outputs and processes. Information Management technology vendors haven’t had that luxury and are providing solutions in a vacuum. These vendors need common approaches that they can leverage within their software.

Open Source

In the Spring 2006 issue of the Financial Services Technology journal “Myths and Realities of Open Source” introduced the reader to the benefits of the Open Source approach to complex code.

The key benefit of Open Source is not, in fact, the free code (there really is no such thing) but the ability to leverage a community of peers who are tackling the same range of problems. The most successful initiatives go beyond program features and functions and offer common approaches to integration, practice and framework.

The best known non-code example of this approach to delivery is the open encyclopedia “Wikipedia” which has rapidly become the first reference of choice for a new generation of

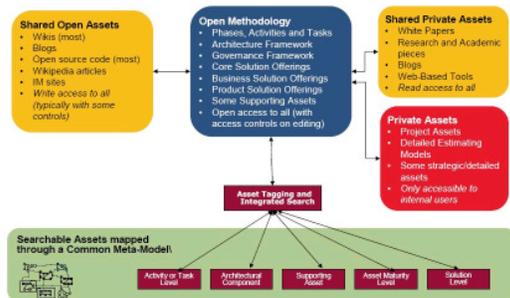
information consumers. Learning from the success of Wikipedia, specialist professional groups are forming to tackle different problem areas as diverse as engineering, law and teaching.

Leveraging the best features of Open Source, an Information Management initiatives can meet the objectives laid out in this article of supporting the methods, techniques, standards and logical architectures necessary to launch the Information Management profession into the 21st Century.

An Enterprise 2.0 approach to standards

One of the most significant differentiators of the Open Source MIKE2.0 approach is that of an Integrated Content Repository, which brings together the open assets from the MIKE2.0 Methodology, shared assets available on the internet and internally held assets.

Any organisation can follow the same approach as BearingPoint and integrate their internally held assets to the virtual hub of assets provided by MIKE2.0 in order to build a community, create common standards, share intellectual property and reduce costs and improve quality through use of reusable assets.



The Integrated Content Repository is a true Enterprise 2.0 solution: it makes use of the collaborative, user-driven content built using Web 2.0 techniques and technologies on the MIKE2.0 site and incorporates it internally into

The Donation of MIKE2.0

In May 2007, BearingPoint officially completed the transition of its core Information Management approaches from a proprietary

capability to an open, shared collaboration under an Open Source “Creative Commons” attribution license). All content is available at www.openmethodology.org and the straightforward license which means that users only need to say where they got the content from, just like a reference in any other published document

Through a six month pilot period, a substantial user base has developed including vendors, Global 2000 businesses and some of the most complex government departments.

While BearingPoint has taken a substantial business risk, losing exclusivity on some of its most valuable Intellectual Property, any loss is more than outweighed by increase in maturity of its clients and the willingness to adopt Information Management principles more widely.

For vendors of Information Management software it provides both a framework they can use within their software and a location where they can describe how to use their solution within the wider Information Management context. In the past they either had to do this function for each of the proprietary approaches or they had to convince the consulting firms to keep up-to-date, now they can invest with confidence in one location.

For Information Management professionals it provides an open forum for the development of their chosen field. With an estimated 50,000 “senior” practitioners world-wide, there has never before been a serious attempt to structure the development of our profession and provide a forum for the most ambitious to promote the development of their ideas.

Finally, and most importantly, for business there is a recognised, non-commercial, framework for Information Management which can be adopted to support all layers in the organization. While it isn’t a silver bullet, it is the first step to knowing what you know, knowing what you don’t know and preparing for what you can’t know before it all goes wrong.

For more information and to contribute, go to www.openmethodology.org.

ⁱ Clayton M. Christensen: “The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail”, Harvard Business School Publishing (1997)

ⁱⁱ Samuel Taylor Coleridge: “The Rime of the Ancient Mariner” (c1798)

ⁱⁱⁱ E. F. Codd: “A Relational Model of Data for Large Shared Data Banks,” CACM 13, No. 6 (June 1970)

^{iv} S. Milgram: “The Small-World Problem,” Psychology Today 1, 60-67 (1967)