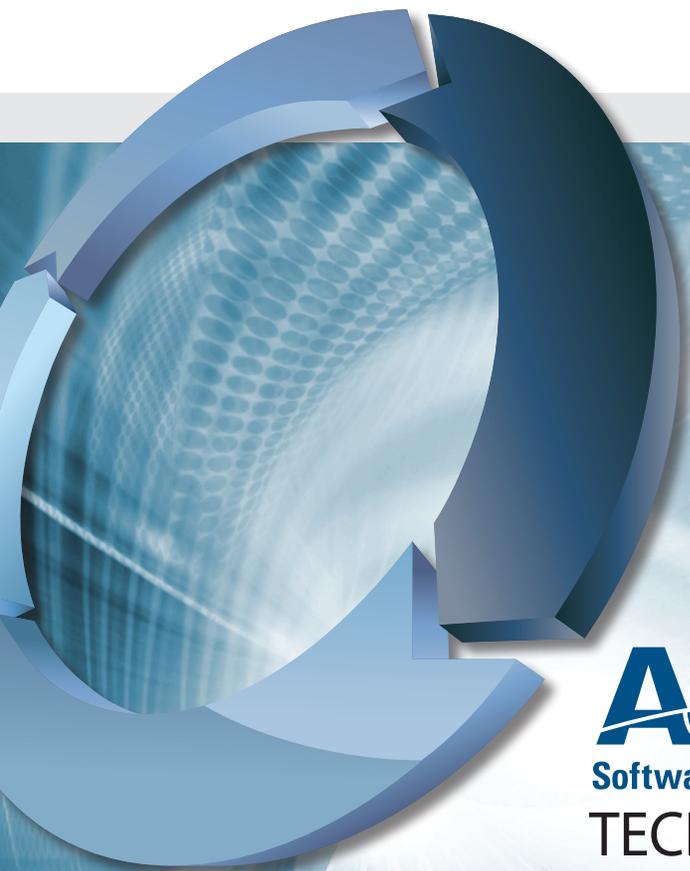


Strategic and Tactical Issues In Metadata Management

by Ian Rowlands



ASG[®]

Software Solutions

TECHNOLOGY TO RELY ON



Strategy without tactics is the slowest route to victory. Tactics without strategy is the noise before defeat.

Sun Tzu: The Art of War

Introduction

We originally consolidated thoughts about metadata strategy and tactics as long ago as 2004. The fundamental issues remain largely unchanged, but there have been enough significant changes in the enterprise IT environment to warrant an update. Some of the striking trends are:

- Shift to knowledge work, expanding the scope of metadata management and driving demand for “easier” ways to consume metadata
- Demand for increased business agility, driving “big data”: and related initiatives
- Increased IT driven business risk driving governance initiatives and the demand for information traceability
- Demand for changing IT cost model exemplified by the emergence of virtual and cloud infrastructure

Effective metadata management requires separating objective and strategy setting responsibility from tactics and implementation responsibility.

Many major enterprises have invested large amounts of money in metadata management projects and reaped significant benefits. Many others have invested similar – or even greater – amounts, and reaped no benefit whatsoever. We argue that the difference is not, necessarily, to be found either in the choice of technologies or in the technical competence of those charged with implementation. What makes the difference is the presence, or absence, of a governance structure for metadata initiatives and the presence, or absence, of a proper separation of strategic and tactical issues. Metadata is an enterprise resource. To allow the same people responsible for implementing metadata management to determine the objectives, strategy and tactics appropriate to the task is to take too light a view of a key source of enterprise advantage. To fail to understand which decisions are strategic and which, though important, are tactical may be to weight some factors too highly and others too lowly.

In this white paper we seek to outline our view of an appropriate governance structure and delineate some strategic and tactical issues.

Elements of the Governance Structure

The governance structure has three elements; charter, organization, and process.

The Metadata Management Charter

Metadata management is an infrastructure issue. It is an investment. There is certainly a return on that investment, but the organization needs to be committed to the proposition that the money spent in infrastructure management could not be better spent elsewhere. It follows that the executive management of an organizational unit in which metadata management is deployed must sponsor it – which implies that if metadata management is being addressed at a truly enterprise level sponsorship must come from the CEO, COO or equivalent. Metadata management must be institutionalized, and not regarded as a discretionary activity.

The Governance Organization

The Governance Organization represents the stakeholders. It should, ideally be independent of the management chain of those charged with implementation. It will likely include:

- A representative of the executive sponsor
- A representative of any department which may use metadata (This carries an implication that, in a true enterprise implementation, every information consuming department should be represented, and metadata governance is likely an aspect of a broader information governance function.)
- A representative of the IT Technology Architecture function

The Governance Process

The Governance Process is followed to determine objectives, set strategy, and validate that tactics and execution support objectives and strategy.

It is not our intention to be overly prescriptive about the detail of the process. It is suggested, however, that there be two regular review cycles – one for objectives and strategy, and one for tactics. The strategic cycle is closely related to the cycle of enterprise strategy reviews (perhaps annual), and the tactical review cycle is linked to specific metadata management projects or is arbitrary and relatively short (perhaps quarterly).

Make sure the distinctions between objectives, strategy and tactics are clearly defined and maintained.



What we mean by Objectives, Strategy and Tactics

Superficially, the enterprise's objectives for metadata management are trivially described. As information is an enterprise asset, and metadata describes information assets, the objective of metadata management must be to ensure that the people in the organization know what information is available, can access it and can use it. Unfortunately, as with many "high-level" definitions, this definition is accurate without being useful.

The objective needs to be:

- Refined to describe specific, measurable objectives
- Tuned to reflect current enterprise goals (essentially to identify which information assets are of most interest)
- Assigned a completion (or at least review) date

Without such refinement it is difficult to determine if metadata management objectives are being achieved!

Ordinarily we would assume that the terms "strategy" and "tactics" were well understood. However, it appears that like most assumptions, this one is poorly founded. We have seen too many strategy documents that contain more words than the tax code. That doesn't mean they are bad documents – but it does suggest that they cover tactics and not strategy, or that they are all inclusive. It may also be worth clarifying that the strategy we have in mind here is not enterprise strategy – it is *metadata management strategy*. The fundamental questions are:

- How, in general, should resources be allocated and organized to achieve metadata management objectives?
- What framework should be provided for the management of metadata, such that metadata management activities can respond to organizational and environmental change in an efficient manner?

Tactics are *metadata management* tactics. The fundamental question is "How, in particular, should resources be deployed to achieve short-term metadata management goals?" Tactics can be invalidated by organizational or environmental change.

Metadata Management Objectives

As before, a reasonable global objective might be "to ensure that the people in the organization know what information is available, can access it and can use it". Some refinement is certainly possible, however. To begin – are we saying that everyone should have access to all metadata? Perhaps not; perhaps there are information assets whose detail may be restricted. Another issue – are we saying that cost does not matter? Does everyone have to have all the metadata no matter what it costs? A good redefinition of the global objective might be "to ensure that everyone has access to all the metadata appropriate to their function, so long as the cost of providing access does not exceed the cost of ignorance."

Metadata Management Objectives should address: What consumers? Which metadata? What business need? How soon? What locations? Which tools?

As well as global objectives, of course, there need to be "local" objectives – objectives for each cycle of delivery. We would suggest that each delivery cycle needs a very simple objective statement – "To deliver to a given group of people, access to a given set of metadata to support a given business purpose. The delivery to be in place by a specified time, using a specific technology to deliver information to defined locations". Of course the general purpose words in the definition need to be made specific for each project, but answering the "who, what, why, when, how, where" questions will clarify objectives and provide for measurability of achievement.

Strategic Issues

It might seem that metadata management strategy should vary from enterprise to enterprise. We are not sure that such is the case. Our experience in selling enterprise repository solutions to more than 700 companies, and implementing metadata management solutions in partnership with all of them has given us some very clear "lessons learnt" that we wish to share.

Scope

Scope defines the extent of the metadata consumer base and the range of the metadata to be delivered. It allows an answer to the questions: "Should we provide access to this group of users?" and "Should we deliver this set of metadata?" Typically it would be described in terms of organizational scope (internal/external, enterprise / division / department / business unit).

The scope of the overall metadata program is becoming very great. As more and more workers become knowledge workers, dependent of the successful exploitation of delivery carry out their work, so the range of information assets expands. Traditionally, metadata management has been concerned with structured operational data and the related models and tools. Increasingly the so-called "unstructured data" or "content" is added to the scope, and most recently information from social media is being added to the mix! Strategically, scope can be very broad ... as long as governance limits the scope of individual projects!

Delivery Cycles

The single most important strategic issue to deal with is how to balance organizational expectations and delivery cycles. There really is a right and wrong way to do this. A recurrent theme in those enterprise repository projects that we have seen fail has been an inability to recognize the enterprise value horizon and ensure that metadata management initiatives deliver identifiable value within that horizon. Put in plain English; if you don't deliver value fast enough, metadata management will lose enterprise support.



Delivery Sequence

The order in which metadata challenges are addressed is a key issue. Superficially, it ought to be possible to rank classes of metadata according to usefulness, and then implement support for them in sequence until the incremental value exceeds the incremental costs. In practice, of course, there will be issues about available resources that constrain the sequence. It is however possible to conceive of an evaluation framework. A number of popular frameworks have been proposed for the definition of enterprise architectures. The most familiar, perhaps, is that defined by John Zachman in his seminal paper "A Framework for Information Systems Architecture."¹

Use an Enterprise or Data Architecture Framework to manage the delivery sequence.

Other EA frameworks have been defined, and Gartner Group has published a useful metadata framework that defines nine metadata sectors. Perhaps we risk some controversy, but we would say that any classification scheme that allows metadata sources to be separated into manageable fractions (to which value and cost can be assigned) will do for the purpose of sequencing. Define a table in which each cell represents a metadata fraction – a manageable fraction being one or more discrete metadata sources which, together, can be supported within the enterprise value horizon. In each cell, record information such as:

- User community – location, size, activity, IT literacy
- Metadata value
- Metadata source
- Cost of metadata capture
- Cost of metadata storage
- Processing requirements (broad brush)
- Any known constraints (including political sensitivities)

Use the table to prioritize cells. After each cell has been dealt with, review the table. Each implementation will change your view of costs, and business changes may change relative values. The actual sequence is tactics. The framework is strategic.

Metadata Topology

One question that straddles the border between strategy and tactics is "where should metadata be located?" The traditional enterprise repository has all metadata transported to a physically centralized location. While this "uber repository" has many benefits, it also has drawbacks in some situations, and alternative strategies are emerging².

Location may be strategic. Currency is usually tactical. Be careful not to confuse the two!

Federated and Distributed Metadata

We define metadata federation as a technique which allows different types of metadata to be stored in and access from different locations on demand through a common management layer, for purposes such as organizational convenience, security or performance. We talk of distributed metadata when the same type of metadata, perhaps sub-divided by organizational unit, is present in more than one location and is not accessed through a common layer. Depending on the motive driving adoption of one or both of these techniques, the choice may be strategic.

"In-place" Metadata

A superficially attractive option is not to transport metadata at all! "Why not", the argument runs, "leave the metadata where it is, and access it directly?" There are, in fact, advantages and disadvantages to this approach. Points to consider include:

- Reduced metadata transport – in principle, only metadata that is required to meet a given information requirement at a given time needs to be transported. This advantage may be illusory. Careful consideration needs to be given to the amount of metadata that may need to be transported at the point of need.
- Elimination of metadata "replicas". Inevitably, moving metadata has transport and storage costs, and introduces the possibility of synchronization problems.
- Lack of a coordinated baseline – if all metadata is left in place, then each metadata store has its own "life cycle". Each metadata store has its own version management facilities (or none at all). This has a specific impact on applications such as data lineage tracking across multiple technologies, as well as posing significant challenges in metadata change management and auditing.
- Challenging model management: Most non-trivial metadata management applications require either a single, semantically consistent metadata model to govern all enterprise metadata and provide metadata coherence or an elaborate process to ensure consistency and integrity.

It should be noted that this approach is not – per-se – an answer to the "active metadata" issue, discussed below as a tactical issue. Active solutions can be delivered from any topology.



A hybrid approach

As well as “pure play” approaches – “Uber repository”, federation, distribution, and so on, there is a more sophisticated option. This might be described as a “use-case-driven hybrid”. There are circumstances in which leaving metadata in silos may be appropriate. Others in which it may be appropriate to build a “corporate” metadata layer from multiple underlying metadata stores. This approach will be especially appropriate for enterprises prepared to invest in a technology framework capable of supporting a variety of approaches.

Tactical Issues

Not all important issues are strategic. There are many issues that have been described as strategic that are, at best tactical. Upon close inspection the error usually arises because metadata management strategy has been commingled with implementation “strategy”. (It is rare that any implementation decision in the Information Technology arena is truly strategic).

Industrialize the metadata management implementation process! Each project is not a “one off.”

Metadata management Project Structure

As implied above, there is a gross structure to metadata management implementation First project, second project, third project There is also a reasonably standard structure to each project:

- Define the target metamodel – an existing metamodel may well support the requirements of the user group you are adding, or the new metadata being incorporated. Sometimes an existing metamodel can be extended. Sometimes it will be necessary to define a new metamodel.
- Implement metadata interfaces – this may be no more than starting to run a vendor-provided function, or it may require special development to be done. Once a collector or interface is created, it will need to execute in a timely fashion.
- Define and implement necessary functionality – the requirements of a given group of users may be met by simple browse and query facilities – or there may be complex “application style” support to be provided. A notable development in recent years, as metadata use has pervaded the enterprise, is the need to embed metadata access into business workflows. As well as changing the required style

of user interface, this requires consideration of workflow and event management capabilities as a required part of the metadata management solution.

- Define the user interface – the user community spans a spectrum, from the metadata management specialist, to the business end-user who really does not wish to know that metadata exists! An understanding of where a particular group fits into the spectrum will affect whether a feature-rich interface is provided, or a web-browser plug-in, or an interface from a business application ... one size will not fit all! It has best practice to define metadata applications “outside-in”, to consider first how the application presents itself to the user, and work in from that point to determine the supporting capabilities, model and metadata.
- Roll out the functionality – without implying that there has been no need for communication with the user community until this point, we do assert that failing to educate the user community in *exactly* how they are to use the new metadata available to them is one of the key missteps in failed metadata management implementations. There is a need to understand the use cases in great detail, and to work through each scenario to show how it works. And then provide a “help desk” facility to make sure that access to the metadata becomes institutionalized, and valuable.
- Monitor usage – the collection of usage metrics allows a proper value to be placed on metadata management activities as well as the identification of problem areas. Some metrics to consider are numbers of customers enabled, number of unauthorized access prevented, number of information assets described, number of accesses to metadata (this may be further refined to count uses of specific reports or queries), instances of erroneous metadata identified and cost of metadata maintenance and access.

Recognizing this standard process and planning around it will reduce the costs of metadata management deployment and accelerate delivery.

The Metadata Management Technical Architecture

It can be argued that, for management purposes, metadata management technical architecture is trivial. There are three layers – metadata creation, metadata management, and metadata deployment.



Figure 1: Generic Metadata Management Technical Architecture

Extraordinarily, many organizations have invested heavily in metadata management without ever explicitly defining this architecture. The definition is critical for cost management and for vendor management. Explicit vendor management provides a key understanding of one of the constraints in developing and supporting metadata management facilities. You need to know each change that will occur in each technology in the metadata creation and deployment layers so that you can work with metadata management vendors to ensure timely support. All too often this is an ad-hoc process.

Metadata management tools

It may seem strange that a vendor of metadata management tools would describe this as a tactical decision. Honesty compels us. There are successful enterprises using a wide variety of technologies. There are three key understandings:

1. There are different appropriate technologies for different requirements – some requirements will be met by tool-specific repositories, some will require an enterprise scope; some requirements will be met by browse and search capabilities, some will require sophisticated underlying processing. Understanding which is which can prevent very expensive mistakes.
2. Different tools (or better said, different metadata platforms) are appropriate at different stages of information management maturity. Typically, we have seen the following stages:
 - a. Documenting: Use of spreadsheets to inventory IT assets
 - b. Exploiting: Use of simple database technology to describe IT assets to support basic use cases
 - c. Enabling: Use of sophisticated tool-specific repositories, sometimes with simple bridges

- d. Pervasive: Use of an enterprise metadata platform to support a wide range of use cases, capable of adaptation and extension
3. Tools do what they do, vendors do not necessarily know your requirements, and will not automatically tell you what their tools do not do! Therefore, if you have requirements that you perceive as critical, make sure that a tool you purchase meets them.

Minimize the number of metadata management vendors you deal with. It reduces management overhead and increases negotiating power.

Tool choice may be tactical. Vendor selection is not. There are enough challenges in integrating metadata. Don't add to them by having too many vendors to deal with. A good choice will be extensible, capable of supporting a range of applications, and not overly focused on one application area.

“Active” or “Passive” metadata

There are two aspects to active metadata:

- Active metadata changes automatically when what it describes changes.
- Active metadata is automatically available and current to its consumers.

Alternatively:

- When what is described by passive metadata is changed, and non-automatic process is needed to update the metadata.
- When passive metadata changes a process needs to be invoked to make the new metadata available and current to consumers.



“Active” vs. “Passive” is usually a financial decision. What is “near enough” to real time, and how much does it cost?

Some metadata management applications will require active metadata – others will not. There may even be applications that are better served by having “fixed point in time” views. There are many applications which do not care greatly – and the choice will be driven by cost. In fact there needs to be a consideration of just what “currency” means in each circumstance.

Effective management of active and passive metadata requires a coherent, semantically correct metadata model that accurately reflects the enterprise’s information and data needs and assets.

Staffing

Plainly there are issues related to staffing a metadata management issues. Should the staff be centralized? Should they be collocated with metadata producers, or consumers, or where should they sit ...? Frankly, we don’t think it matters very much. There are, however, two things that matter a great deal: All the roles need to be covered, and there needs to be an explicit plan to provide the appropriate skill sets.

Metadata management is not a trivial activity. Would you staff your RDBMS environment without investing in training the staff?

Clearly, both the size of staff, and the appropriate skill sets will vary as a function of the technologies in use. We have commonly found these roles:

- Executive sponsorship
- Metadata team management
- Metadata strategic planning
- Metadata tool administration
- Metadata modeling
- Project Management
- Development (often splitting into requirements analysis, design, functional implementation and user interface implementation)
- Infrastructure Management (As strategic systems move from a mainframe environment to a web-based environment, it is increasingly the case that the complex interactions are not fully understood.

Although it is not really a staffing issue, it also makes sense to consider roles in the user community. Plainly these are much more application-dependent, but we have commonly found three categories – “metadata expert”, “power user” and “consumer”. The three categories have different requirements for access capabilities. Once roles have been identified they can usefully be mapped against facilities required using an authorization matrix.

Conclusion

Enterprise metadata management as a competitive advantage is becoming an important strategic weapon in today’s evolving business environment. Early adopters of enterprise metadata management will have the lead with experience, trained professionals, installed base, and the subsequent advantage. The issues discussed in this paper are not new. They are the same issues faced when developing, implementing and adapting any new methodology or technology. The major difference here is that effective metadata management, if carried to the full extent, can add value across all areas of the enterprise from business management to technology implementation. The challenge is to develop strategy and tactics which best supports the enterprise, and then to stay the course of implementation and management.

Sources

¹ Originally published in IBM Systems Journal, Volume 26, Number 3, 1987. Reprinted Ibid. Volume 38 numbers 2 & 3, 1999, and in other places.

² Ibid



A recognized innovator in enterprise IT software solutions, ASG Software Solutions has been optimizing 85 percent of the world's most complex IT organizations for more than 25 years. We create and deploy unique software solutions that reduce cost, mitigate risk and improve service delivery throughout the IT lifecycle. ASG's comprehensive solutions help you solve today's challenges, such as cloud computing and big data, while driving your business forward by providing insight and control across cloud, distributed and mainframe environments. ASG is a privately held global company based in Naples, Florida, with more than 70 offices worldwide.

www.asg.com

ASG Worldwide Headquarters | 1.239.435.2200 or 1.800.932.5536
1333 Third Avenue South Naples, Florida USA 34102

Copyright © 2012 Allen Systems Group, Inc.

All products mentioned are trademarks or registered trademarks of their respective holders.

ASG-Strategic_and_Tactical_Issues_in_Metadata_Management_white_paper_20111207en