



Unlocking information and the Internet

White Paper

Version: 1.0

15 May 2001

Keith Sherringham

Copyright © IMS Corp. All rights reserved.

IMS Corp.
Asia - Pacific Headquarters
Suite 5 275 Maroubra Road
Maroubra
NSW 2035
Australia

Tel: +61 (0)2 9314 2908

Fax: +61 (0)2 9314 2908

Email: info@imscorp.com.au

Web: www.imscorp.com.au

© Copyright IMS Corp. All rights reserved.

“to and from any device, anywhere, any time” is a pending trademark of IMS Corp. All other referenced trademarks are those of their respective owners.

This document is for information purposes only and any advice given is of a general nature only and may not be applicable to an individual or a specific situation. IMS Corp. accepts no responsibility for any consequential loss or damage arising from the use of this document.

Abstract

The key to lower costs, to accessing corporate knowledge and the power of the Internet is in the systematic application and use of metadata (data about data). Metadata allows you to find information, it tells you what the information is, you know who has access to that information and more. Through the systematic application of metadata to information at source, by the subject matter experts, as part of an everyday business process, not only is corporate knowledge and corporate expertise trapped and made accessible but also considerable gains in productivity and cost savings realised. With use of the Internet and knowledge worker optimisation, it is no longer a question of Return On Investment from metadata but a question of gross business negligence if metadata is not done.

1 Introduction

The key to lowering costs, to accessing corporate knowledge and to harnessing the power of the Internet is in the systematic application and use of metadata.

Corporate knowledge and corporate expertise are considered to be one of the most valuable assets a company has and is often the key IP a company has. This is how a company differentiates itself in the market. Access to and management of this corporate knowledge and expertise however, continues to be a challenge faced by business. The often overlooked but simple use of metadata (data about data), maintained at source, by the subject matter experts, is one of the most effective ways of trapping, managing and sharing this asset.

2 Benefits of metadata

The main benefits of metadata are:

Access to information - By being able to quickly find information, time is not wasted nor unnecessary effort spent on duplication. Re-invention is no longer required.

Right information – With metadata telling you this is the latest version of the document, when it was last changed, who changed etc. You know you are using the right information.

Security – Metadata tells you who can access information and what they can do with it.

Propels business optimisation – By having people adopt metadata as part of an everyday business process, you start to propel the business optimisation necessary and begin the processes of information and knowledge management.

Knowledge worker optimisation – The Internet is optimising the role of the knowledge worker with a resultant change in focus to high-end value-added services. This can not be achieved without access to information and it is metadata that ensures this.

3 The role of the Internet in business

The Internet has not changed the challenges a business faces. What it has done is change the way the issues need to be addressed. The issues of customer service, time of delivery, quality assurance, lower cost and expectation management etc. are still the same. The Internet has however, provided more channels to be supported, has increased the pace of change and caused customers to expect a higher level of service at no extra cost. The Internet also draws an inefficient business process or shortcoming into sharp focus, meaning that it can no longer be covered up. Online selling may vary



the number of sales made within a particular channel but a total increase in sales comes from a whole of business approach.

The Internet is about lowering costs. Part of this cost saving is optimising the role of the knowledge worker. Knowledge workers access data from a range of data repositories or silos, pull the information together and present it back to a customer, e.g. Customer Service Representatives in call centres, account query or order tracking, Figure 1.

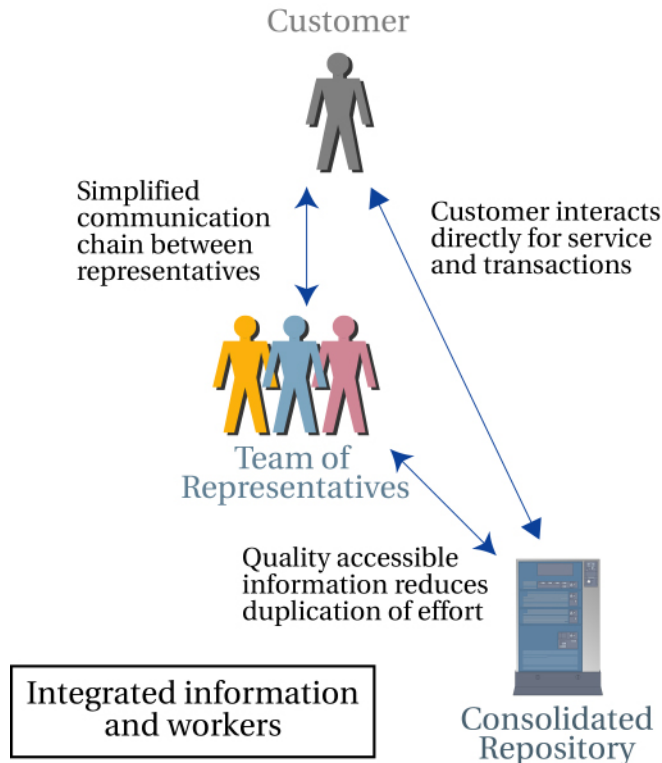


Figure 1. Accessible information and work-flow lowers hidden costs and ensures better customer service.

4 Audience – Task – Recipe

Every day of the week within business, products are sold, billing issues resolved, customer service provided and more. You know who your customer is (Audience), you find out what they want to do (Task) and you walk them through a series of steps to accomplish that task (Recipe). Whether writing a speech for the CEO, manufacturing a car, selling a product (Figure 2) or solving issues in a call centre, employees are familiar with the Audience – Task – Recipe (A-T-R) approach because they do it every day of the week.

The A-T-R approach is a shared experience for both internal and external facing audiences and every business activity can be defined accordingly¹. The importance of the A-T-R concept is seen when business processes are brought to the Internet and in the management of information.

¹ This is the entire bases of assembly lines and manufacturing, the issue is its application to services.



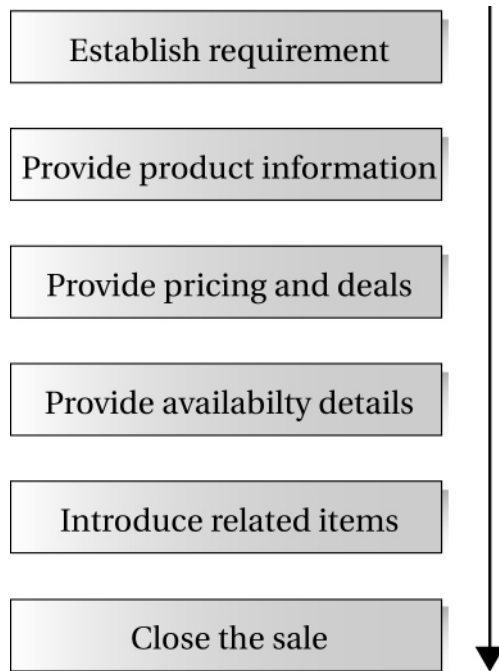


Figure 2. The sales process in its rawest form.

5 What is meta data?

Metadata is a term that is often misused and/or carries a negative association. Metadata is simply data about data. Consider a payment for services. The data is the amount paid but all of the other related details like who paid it and when it was paid is metadata. With a legal contract, the contract itself is the data but the date it was signed, who the contract was between, where the file is stored, the format of storage, that it is a contract and who has access to the contract is all metadata.

Thus, it can be seen that metadata is important because it lets you find the information you need, it tells you if it is what you are looking for, whether it is the correct item, who should look at it etc. Ready access to all of these details is the bases of a smooth running operation.

Every content item, e.g. billing record or document, has its own associated metadata. The metadata describes the content item. A content item may be a collection of other items (parts) that have been brought together. Consider the case of all of the information within an organisation. If this was considered as one entity, it can be said that this is owned by the organisation, on this date with this monetary value. The data are all of the organisation's information and the metadata is owner, date and value.

Hence the amount of and type of metadata vary not only with the different data held in an organisation but also in how it is considered and applied.

6 Role in business

Before considering metadata in detail and the value-add it brings to business, its position in business needs to be established, Figure 3. This simple model shows the corporate repository at the base which holds all of the information within an organisation, e.g. billing records, financial reports, documents etc. To access this information you need to know that it exists and what it is - metadata. The information is then delivered to a range of users through presentation management. Any search capability must then be able to search the metadata as well as the information in the repository.



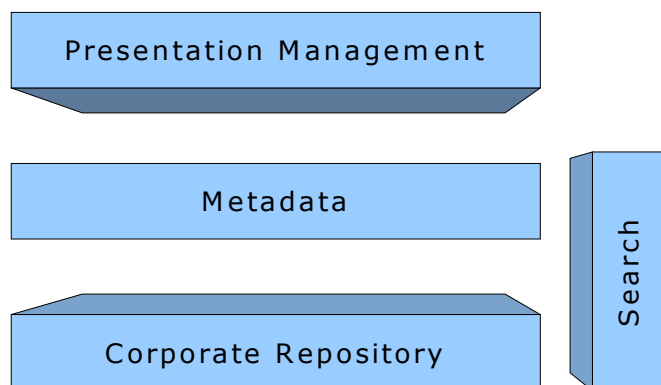


Figure 3. Position of metadata in business.

7 Realising the investment

To realise the benefits of metadata a whole of business commitment is required. As with other business activities, it takes time, effort and resources to do metadata activities and it needs to be managed. If metadata is not seen as every ones responsibility and is not done as part of an everyday business activity, then the activity will yield little Return on Investment (ROI). Metadata is only as useful as the quality of information provided – this determines the ROI.

The problems to avoid with metadata are now outlined:

Many metadata activities have failed because the information was maintained centrally by specialists and done as an after thought to other business processes. By having the subject matter experts do it at source, as part of an everyday business activity, these problems are mitigated.

As metadata helps you find information, the natural tendency is to want more of it. However, someone has to supply and maintain the information and herein lies the problem. Having a small number of populated and well-maintained metadata fields is of more benefit than having many fields sparsely populated that are out of date.

Making the process of metadata entry easy facilitates the process. Use of clearly defined fields, default values customised to the user, auto-population of fields, system values like dates that are automatically adjusted and short pick-lists that are understood, all facilitate the process. Use schema that users understand and are intuitive (business process driven rather than classification driven).

When asking people to do additional work or something that they are not used to, it has to be easier and of benefit. They need to have it explained to them, they need to understand the advantages, they need to be trained and they need to see the benefits on a daily basis. Incentives for doing metadata need to exist. Likewise, disincentives for not doing metadata and for poor workmanship should also exist.

Doing metadata properly from this time forward and addressing the backlog as business priorities dictate avoids that sense of desperation and being overwhelmed.

8 Types of met adata

As mentioned previously, a range of metadata can exist, varying with data types, industry and playing different roles within business. However, some common metadata also exists, as is the case



with documents. Much of the information within an organisation exists as documents and the following common information is of use when searching and using documents, Figure 4.

Element	Description
Title	The title of the content item
Abstract	A short informative summary or precis of what the content item is about
Creator	Name of the person who created the content item
Contact	Name of the person to contact about the content item
Date Created	Date the content item was created
Last Edited	Date the content item was last edited
Edited By	Name of the person who last edited the content item
Review Date	Date at which the content item comes up for review

Figure 4. Sample document metadata.

8.1 Content types

Classification of information into content types is part of the metadata required. This is done on a routine basis within business and is common across business, e.g. creating a budget or writing a tender, Figure 5. Trapping this information at source and making it available to others can be easily achieved.

agreement	glossary	purchase order
bibliogrphay	guide	rates
brief	help	report
brochure	invoice	request
budget	manual	requirement
card	memo	requisition
catalogue	news	roster
chart	plan	schedule
conditions	portfolio	specification
contract	policy	statement
correspondence	presentation	tender
directory	procedure	terms
docket	profile	ticket
form	promotion	timetable

Figure 5. Sample content types.

For transactional data, e.g. a billing record, this information is trapped at source by the applications used. It is the flexibility of being able to retrieve all memos written between two set dates, which is required.



8.2 Access permission

Knowing who has what access rights to what information is all-important. Access rights need to exist within and between organisations. Who within an organisation has permission to view a document or a billing record and who can change it? Who can edit a payment record within an organisation after payment but who can apply the payment?

Whilst the metadata associated with transactional data is set by the system and the login rights, that for documents is often unknown. By having the creators of the document – the subject matter experts – allocate permission at the time of creation, people know who can see that document and who can edit it etc.

Associated with access rights is that of useability / purpose / suitability for use. Whilst someone may have access to information, this may or may not be the same as being able to understand it and to use it. This is the basis of the Audience – Task – Recipe approach. This is the real business value-add.

Access permission exists by group, individual and organisation.

8.3 Language metadata

One component of the metadata that is often overlooked is the language in which the information is stored. For businesses that operate on a global basis, multi-lingual support is a standard business practice. By language versioning documents from a base source and use of metadata, multi-lingual information can be easily managed.

8.4 Device metadata

The ability to deliver transactional data to and from, any device, anywhere, any time has existed within organisations for many years. The challenge brought by the Internet is to deliver documents and transactional data to any device, irrespective of organisation. Whilst a 100 page tender as a Microsoft Word document may be delivered to a Web site, sent via e-mail and FTP to a desk top or rendered on a handheld device, it is unlikely that it will be displayed on the small screen of a mobile phone.

Through the systematic application of metadata on mechanism of delivery, people know which documents can be sent to whom by what device. As business becomes more automated, this information will become critical – information will not be able to be delivered without. Device specific metadata will also need to exist.

8.5 Keywords

One of the issues with searching and metadata has been the use of keyword classification schemes. Metadata activities usually involved creating lists of keywords to categorise information and then people were expected to use these extensive lists and to systematically apply them on a consistent basis. Taxonomic experts have been able to do this for flora and fauna, selected data directories exist and whilst librarians have been able to do this on a consistent basis for collections of books, there are very few other examples of this working effectively. These examples work because specialists spend all of their time doing it, something that does not exist within daily business for information management.

Traditional keyword lists shall continue to play a role in information management and are still valid for certain types of data and applications but new dimensions are needed to support business information, particularly documents, as the Internet continues to optimise the role of the knowledge worker.



8.6 Classification schemes

As discussed previously, business operates using a series of proven processes, customised for a specific audience: Audience – Task – Recipe (A-T-R). In contrast with conventional knowledge classification schema (see previous section), business people are familiar with business processes, they interact and operate in this way and require the right information at each step of the process to complete the job. **Information is classified by purpose according to the A-T-R approach.**

Audience information is established partly through the access permission but also by the subject matter experts at time of data entry. Organisation criteria are granted in the access permission and at the time of data entry, the subject matter experts selects the role, e.g. network administrator, project manager. This classification is a natural extension of the role-based access that is present within business.

Task information is partly set by the role determined and partly from other sources. The subject matter experts at the time of data entry allocate the content item to the task and at the required step. That is your classification list is a list of tasks and steps by audience.

A piece of content may be assigned to multiple tasks and multiple steps within a task but the amount of duplication is actually quite small. A Network Administrator does not need to do corporate procurement or analyse marketing material. If they do, they login as these roles, the information classification does not change.

Having to select multiple organisations and multiple roles at time of data entry may seem excessive but consider the following:

- This is often done already but it is not formalised and people spend time searching for this information from other sources.
- The access rights have to be set anyway, it is merely extending existing information and applying it to purpose / useability / suitability.
- It is a case of setting exclusions, not inclusions.
- Corporate knowledge and expertise is trapped at source, saving on costs, time and duplication of effort.

If the A-T-R approach to classification is not properly implemented and planned then its benefits will not be seen and may be even be counterproductive. This is true of any schema however. As with other classification schemes, one starts simple and develops as the complexity grows.

After a while, the patterns of information associated with the A-T-R approach will become evident and pattern-matching software can be used to assist with metadata population. Existing classification schemes do not blend themselves to automated processing by pattern matching software.

8.7 Related items

Another way in which corporate expertise is trapped in metadata is through the use of associated links or related items. If a legal officer reviews a contract for a selected company, then proven expertise shows that they may also be interested in certain related items or completing certain tasks. When this is repeated for specific audiences by selected tasks, the corporate expertise is trapped and becomes accessible to others.

Incidentally, this use of related items provides effective Web site navigation, forms the bases of meaningful personalisation and further reduces the need for complex classification schemes that business people are unfamiliar with and reluctant to use.



8.8 Content item application

The discussion on metadata has mentioned that it describes the data and the need to apply it at the item level. This may seem daunting at first but when the following are considered, this is not the case:

- For transactional data, the permission is set automatically by the system. The information is automatically selected at the right step, in the right task, by the right audience.
- For documents, the need for manual metadata population is more significant but it is needed anyway and the duplication between tasks and steps is small.
- The business need for these activities already exists; the Internet has merely increased the pressure to do it.
- Free text searches etc. can still be used to find information by all audiences. The A-T-R approach merely makes it quicker and easier.

9 Metadata and presentation management

The way information is presented to users, whether in printed matter or on Web sites, is presentation management. The management of information is a back office process done on a daily basis as a whole of business activity. The automatic delivery of content to and expiry from Web sites, is presentation management. Once this difference is recognised, the importance of metadata in presentation management can be seen.

With metadata telling you the information version, its currency and more, content can be automatically displayed or removed from a Web site. It also ensures that the right information is used in a marketing brochure. With adoption of the A-T-R classification scheme, Web site navigation can be automatically determined and optimised from the data, realigning the role of Web site managers to high-end value added services.

With information about the information, audiences, related items and use patterns being trapped at source, higher quality marketing material can be produced at less cost and focused accordingly. Having this corporate experience trapped, allows for further automation in the presentation and delivery of information via electronic media.

10 Search

As shown in Figure 3, the search capability needs to access both the metadata and the information directly. The key to successful searching is the ability to resolve the problem of "I don't know what I really want but find it anyway!" That is a user sort of knows what they want but has difficulty articulating it and is not sure of the search criteria. This is hard to resolve using traditional search approaches because the classification scheme is unfamiliar to the user. When information is classified according to common business processes however, the chances of the right information being found are higher because the classification process is familiar.

Existing search engine technology can work with or without the use of classification schemes but they can make it more efficient. The use of the A-T-R approach for classification does not decrease the capability of the search software; it can actually reduce the demands made of search engines because information is easier to find. Just as some search engines optimise around a classification scheme, e.g. taxonomy, so the case may be with the A-T-R classification approach.



11 Business value-add

A recurring theme within the discussion has been the application of metadata, by subject matter experts, at source, as part of an every day business process. Whilst this ensures that metadata is populated and corporate knowledge is trapped, this may not necessarily equate to what is the best business value for the information. A network administration would tend to classify according to network values, whilst a sales representative may value that same information in a different way. This issue of value or perspective has important impacts upon the way metadata is applied and used.

Firstly, it become evident that not one person can put in all of the metadata for a content item, it will be a shared responsibility.

Secondly, the same information may have several classifications attached to it, according to the different audiences. Someone else within an organisation may metadata the same item from a different perspective and “this versioning” needs to be tracked and managed. The implications of this in terms of business processes and the current perceived roles and responsibilities are significant but these issues exist with any business and classification scheme.

The A-T-R approach places an emphasis on business value. This is a positive business driver that is absent from existing classification schemes. As the use of metadata grows and the business value is applied more, an additional value to the business of metadata will occur. The on-sale value of information within many businesses is not known, particularly certain combinations of information for selected audience. With this information being recorded and being used, an additional dimension to the sales and marketing process is realised.

Finally, by having metadata populated and maintained at source by the subject matter experts, you propel the business optimisation necessary.

12 Summary

The business need for the systematic application of metadata, by subject matter experts, at source, as part of an every day business process has been in existence for years. Those that have done this have been able to lower hidden costs and provide better customer service. The pressure on business to lower hidden costs using the Internet has never been more intense. Optimising the role of the knowledge worker requires quality metadata. The key to unlocking the power of the Internet and for accessing information lies with quality metadata.

For metadata to be successful, it needs to be maintained at source, by the subject matter expert, as part of an everyday business process. Traditional classification schemes can still be used to assist with information retrieval but these are often cumbersome and not understood by the business. By using proven business processes and the Audience – Task - Recipe approach to classification, not only is the process of metadata population and maintenance by business people facilitated but information retrieval and management is also simplified.

