



# CIO Council

\*\*\*\*\*CHIEF INFORMATION OFFICERS COUNCIL\*\*\*\*\*

About The Council

Committees

Documents

Events

Federal Gov't CIOs

Links

Search

What's New

Home

[Text Only](#)

## What Every CIO Needs to Know About Metadata

February, 1999

*This is the first of a series of CIO recommendation papers from a sub-group of the Federal Agencies IT Architecture Working Group, part of the Federal Government CIO Council's Interoperability Committee. Follow-up papers will be addressing specific issues in repositories and metadata management. The recommendations in these papers will give concrete support to various aspects of the government-wide IT architecture conceptual model. This paper speaks to a foundation issue.*

Metadata IS data. It's just plain, ordinary, every-day, garden-variety data. What makes it metadata is how it's used. It's USE that adds the "meta."

The term was coined by folks who didn't want to say "data data." Everyone understands such terms as "payroll data," "personnel data," "inventory data," "medical data," and "budget data." The second word, "data," is the noun, and the first word is a modifier that says what kind of data we're talking about, or how the data is being used.

"\$100.00" is a piece of data. It could be payroll data, personnel data, inventory data, medical data, or budget data. We need a DESCRIPTION to tell us what kind of data it is. That description is data about data, or "data data," otherwise known as metadata. Metadata tells us the meaning and context of the piece of data.

Metadata also tells us how to understand the way the data is expressed. It tells us that "\$100.00" is a monetary amount in U.S. dollars, expressed in terms of dollars and cents. That's how the data is represented. So, metadata tells us both the meaning of the data and how it's expressed or represented.

### The Game Isn't Just Data Any More

For years, the term "metadata" in the information technology (IT) world was used just for data, and just by data administrators and data base administrators. That's why it's associated so strongly in most peoples' minds with data base activities.

Now the world of IT is much bigger than just data and data bases. There's Word Processing. There's E-mail. There's Voice-Mail. There's Image Processing. There's Multimedia. Most of all, there's the World Wide Web.

All these must be described in order to be managed. They must be measured and reported, and described in all significant aspects. This means new terms and new strategies must be brought into play, to talk about each as a technology and the information that's involved in each. We must be able to talk about

- the information in a word processing file as well as the word processing file and system themselves
- the information in an e-mail message as well as the e-mail message and system themselves
- the information in a voice-mail message as well as the voice-mail message and system themselves
- the information in an image as well as the image and imaging system themselves
- the information in a video as well as the video and video system themselves
- the information in Web pages as well as Web sites and their use.

NOW when we talk about metadata we're talking about all the data activities we used to talk about PLUS tons of new information description activities. Data is still playing in the game, but the game now is more about information as found in documents, messages, images, sound streams, and videos. The same word, metadata, applies to the descriptive data for both.

Let's look at some examples:

<u>The Information Package</u>	<u>Part of its Metadata</u>
Video	Subject and participants
Movie	Copyright
Photograph	When and where taken
Engineering drawing	When and by whom made
Book	Date and place of publishing
E-mail message	Date and time of delivery
Voice-mail message	Date and time of recording
CD-ROM	System requirements for use
Image	Format and standard used
Web page	URL

### **Why the Big Deal about Metadata?**

Metadata lets us find information. Imagine a file room with no discernible order in the labels on the cabinets, no discernible order in the labels on the drawers, and no discernible order in the labels on the file folders inside the drawers. Imagine your PC if you couldn't give files your own labels and the system assigned labels with a random number generator. Imagine the World Wide Web with no search engines or search engine possibilities.

Metadata lets us understand important things about information packages. Imagine e-mail messages without names or identifications of senders, without names or identifications of addressees, and without dates of origination. Imagine office memoranda without the names of originators and intended receivers, without subject or reference headings, and without dates of origination.

## **OK. But why does a CIO need to be concerned with metadata?**

Metadata is one of the biggest critical success factors to sharing information. Metadata also is one of the biggest critical success factors to storing information cost-effectively. Metadata can make your information sharing and storage efforts great successes, or great failures. Metadata costs money and has its own ROI. Metadata can get you in trouble with the law, or keep you out of such trouble.

## **If that's the case, How do I manage metadata for success?**

Simply put, metadata management is making metadata do more things for more people in more ways, more economically.

You manage it through people who understand it, and who have and know how to use the tools of metadata management. Foremost among those tools are Metadata Registers, Dictionaries, Directories, Locators, and Search engines.

## **What's the difference between Information Management and Metadata Management?**

Metadata management is a key part of information management. It may be all of information management in situations where information management focuses entirely and exclusively on information sharing. However, for many organizations, information management includes also defining the information needs of people (who needs what information), and designing the information creation or collection activities (the processes of obtaining, maintaining, and using information). These aspects involve metadata, but they also involve much more, including business processes.

## **Where does Metadata Management fit, organizationally?**

If you have an information management organizational unit that really does agency-wide information management work, then metadata management belongs there. If you have such an organizational unit, but the real work of information management is being done elsewhere, put metadata management where the real work is being done. If you don't have an information management organizational unit and information management work is done on a scattershot, every-group-for-itself basis, then you have to start at square one with information management, and look long and hard at where and how information sharing will play in your agency.

## **Is there anything else I need to know?**

There's a legal side to metadata. As business processes are changed from being paper-based to being conducted electronically, such as electronic commerce and electronic filing, the metadata aspects can affect significantly the design of systems and the cost of systems. The metadata that a system creates to support the finding and understanding of information by its users is also the same metadata that third parties rely on to ferret out potentially damaging information. It's also the same metadata that courts and regulatory agencies are declaring must be treated equally with the information being described, and that must be preserved together with the information being described. There is a growing body of law that gives third parties access to the metadata created and used operationally within systems, and that imposes evidentiary obligations for that metadata upon system managers - obligations that they previously had only for the information itself.

## **Bottom-line**

The alternative to metadata management is information chaos.

Send questions or comments to: [The CIO Council](#)