



# TECHNICAL REPORT

## The Enterprise Information Tag (EIT) - Information Management's Missing Link

December 26, 2004

ANSI/AIIM TR47-2004

## Technical Report for Information and Image Management –

---

# The Enterprise Information Tag (EIT) — Information Management's Missing Link

---

a Technical Report prepared by  
AIIM and  
Registered with ANSI

December 26, 2004

### Abstract:

This report urges creation of a technical specification for applying enterprise-determined identifying metadata to any kind of electronic information in any media, such that the metadata can be used to find and manage the information according to enterprise policies and processes. The metadata association is formatted in a standard way for conveyance with the information and for display. This formatted association between information and its metadata would be built into commercial-off-the-shelf software products of all kinds, and would work in conjunction with enterprise-created and managed metadata repositories. Because the tag is so closely tied to enterprise metadata management processes, the concept is a new, integrated infrastructure approach to information management within an enterprise. As such, its direct benefits are in evidentiary management, security management, privacy management, content management, knowledge management, legal and regulatory satisfaction, intellectual property management, and business process management. The authors believe that the vision, when built into commercial products and services, will yield extremely significant economic benefits to all kinds of enterprises in all sectors. Although the vision will take several years to realize fully, it promises a win-win for everyone.

Copyright 2004 by  
AIIM  
1100 Wayne Avenue, Suite 1100  
Silver Spring, MD 20910-5603 USA  
Telephone: 301/587-8202  
Fax: 301/597-2711  
Email: [aiim@aiim.org](mailto:aiim@aiim.org)  
Web site: <http://www.aiim.org>

ISBN 0-89258-406-8

No part of the publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means—electronic, mechanical, photocopying, recording, or otherwise—without prior written permission of the publisher.

Printed in the United States of America.

## Contents

|                                                                                                          |           |
|----------------------------------------------------------------------------------------------------------|-----------|
| <b>1 Purpose and scope .....</b>                                                                         | <b>1</b>  |
| <b>2 Introduction.....</b>                                                                               | <b>1</b>  |
| <b>3 Uniqueness of the concept.....</b>                                                                  | <b>2</b>  |
| <b>4 Acronyms and definitions .....</b>                                                                  | <b>3</b>  |
| <b>5 The concept.....</b>                                                                                | <b>4</b>  |
| 5.1 An infrastructure strategy .....                                                                     | 4         |
| 5.2 Information objects (IOs).....                                                                       | 4         |
| 5.3 Enterprise Information Tag (EIT) .....                                                               | 5         |
| 5.4 Business case .....                                                                                  | 6         |
| 5.5 Evidence .....                                                                                       | 7         |
| <b>6 EITs and enterprise management needs .....</b>                                                      | <b>7</b>  |
| 6.1 EITs and privacy management .....                                                                    | 7         |
| 6.2 EITs and security management .....                                                                   | 8         |
| 6.3 EITs, content management, and knowledge management .....                                             | 8         |
| 6.4 EITs and applications management.....                                                                | 8         |
| 6.5 EITs and infrastructure management.....                                                              | 9         |
| 6.6 EITs and intellectual property management .....                                                      | 9         |
| 6.7 EITs and evidence management .....                                                                   | 9         |
| <b>7 Path to realization.....</b>                                                                        | <b>10</b> |
| <b>Annex A (informative) Illustrative scenario: Monolith Financial and Security Services Company</b>     | <b>11</b> |
| <b>Annex B (normative) Structure, format, and operation of the Enterprise Information Tag (EIT) ....</b> | <b>16</b> |
| <b>Annex C (informative) EIT generation and mechanics, EIT metadata, and metadata management</b>         | <b>19</b> |
| <b>Annex D (informative) Frequently Asked Questions (FAQs).....</b>                                      | <b>25</b> |
| <b>Annex E (informative) Background of this report .....</b>                                             | <b>29</b> |
| <b>Annex F (informative) Bibliography .....</b>                                                          | <b>32</b> |

## Foreword

(This Foreword is not part of ANSI/AIIM TR47-2004—*Technical Report—The Enterprise Information Tag (EIT)*. It is included for information only.)

This technical report was prepared by AIIM Standards Committee C22, Evidentiary Support for Electronic Information. The report provides a concept for automatically tagging electronic information in a standard way to help identify, protect, find, use, and dispose of stored information, and it urges creation of a technical specification for the concept. The formulation of the concept is stated in the normative Annex B.

Five informative annexes are included with this Technical Report: Annex A provides an illustrative scenario of uses and benefits of the Enterprise Information Tags; Annex C addresses EIT generation and mechanics, EIT metadata, and metadata management as they relate to this report; Annex D presents some frequently asked questions (FAQs); Annex E provides background on the actions leading to this report; and Annex F lists a bibliography of publications mentioned in the text.

While this report speaks to several management topics, including security, privacy, intellectual property, and records, it is not a report about or for any one of them. Rather, it is an overarching information management report for all of them. When industry creates the technical specification envisioned herein, the C22 committee anticipates that professionals in each of the subject areas will develop recommendations for their profession's use of the technology.

Suggestions for improvement of this Technical Report are welcome and should be sent to the Chair, AIIM Standards Board, 1100 Wayne Avenue, Silver Spring, Maryland 20910-5699.

At the time it approved this recommended practice, the AIIM Standards Board had the following members:

| <b>Name of representative</b> | <b>Organization represented</b>             |
|-------------------------------|---------------------------------------------|
| Charles Dollar, Chair         | Dollar Consulting                           |
| Leslie Banach                 | IBM Corporation                             |
| Betsy Fanning                 | AIIM                                        |
| Chuck Fay                     | FileNET                                     |
| Bruce Holroyd                 | Eastman Kodak Company                       |
| Virginia Jones                | Newport News Department of Public Utilities |
| Linda Koontz                  | U.S. General Accounting Office              |
| William E. Neale              | U.S. TAG to ISO TC 171                      |
| Dan Schneider                 |                                             |
| Louis E. Sharpe               | Picture Elements Inc.                       |
| Chris Thompson                | Recognition Research Inc.                   |

The AIIM Evidentiary Support Committee, C22, had the following voting participants at the time it processed and approved this report.

| <b>Name of participant</b> | <b>Affiliation</b>                    |
|----------------------------|---------------------------------------|
| Daniel B. Schneider, Chair |                                       |
| Robert Blatt               | Electronic Image Designers, Inc.      |
| John Breeden               | Virginia Department of Transportation |
| Robert Breslawski          | Eastman Kodak Company                 |

**ANSI/AIIM TR 47-2004—TECHNICAL REPORT — THE ENTERPRISE INFORMATION TAG (EIT)**

Michael Cherry  
Rae N. Cogar  
Vigi Gurushanta  
Juris Kelley  
Nancy L. Richards  
Tom J. Sas  
Janet P. Strauss

GMC7, Inc.  
RCS Consulting  
eVIDA Group  
Knowledge in Motion LLC  
City of Bellevue, Washington  
Hewlett-Packard Company  
National Geographic Society



# Recommended Practice — The Enterprise Information Tag (EIT) Information Management's Missing Link

## 1 Purpose and scope

This report calls for the creation of a technical specification for a new capability in commercial software products. The capability will be an enormous help and economic benefit to enterprises in managing their electronic information holdings. The report explains the concept behind the specification and how it will work in practice. The concept is illustrated in a scenario in Annex A and described normatively in Annex B. Additional information is given in Annexes C and D.

The needed specification must be developed by the major software product companies as a consensus national or international standard. Once the standard has been agreed upon and supported in commercial software products, the annual savings and economic benefits to the enterprises using the products are likely to run in the billions of dollars.

This report is being issued by a standards committee of AIIM International but is actually universal and global in its scope. It applies to electronic information holdings of all kinds and in all media, industries, and countries. It applies to all software product vendors, operating systems, and electronic technologies. It is independent of laws and governmental regulations. It offers a new and powerful tool for privacy management, security management, knowledge management, applications management, infrastructure management, intellectual property management, and evidence management.

This report has been many years in the making. Its background is related in Annex E.

## 2 Introduction

Information of all kinds and sources is being held on devices and in storage media owned or controlled by enterprises and their employees. All of that information, from interpersonal communications to data warehouses, is both a potential asset and a potential liability. Government regulations are prescribing what information must be kept electronically, and even how it must be kept, while court cases are showing how vulnerable all information holdings are to discovery in lawsuits. Enterprises are searching for ways to describe their information more productively, to get more value from it, and to manage it more cost-effectively, while their adversaries are using enterprise information management weaknesses to create discovery burdens and get discovery findings that can favor plaintiffs in pre-trial settlements and jury awards.

Information holdings and their management impact many aspects of enterprise operations besides evidence of deeds and possible misdeeds. The management of the holdings is crucial to effective privacy management, security management, intellectual property management, applications management, and even the management of IT infrastructure. The simple fact of today's world is that information is dollars to the bottom-line. Not just the cost of acquiring and processing it—that's the least of it—but what it can return to the enterprise if it's the right information, at the right time, in the right place and used the right way, and what it can cost the enterprise if it violates a law, infringes on some legal right, leads to harmful action, or otherwise violates enterprise policy.

In 1995, several AIIM member user enterprises united behind the idea of creating a standard way of automatically tagging electronic information with metadata that would help identify and locate information, help protect and use the stored information needing retention, and help eliminate stored information according to enterprise policies. The concept was called initially "REBI," for Reliability of Electronic Business Information. Many representatives of major corporations and government agencies endorsed it. The AIIM Standards Board chartered it as a project of Standards Committee C22, Evidentiary Support. (See Annex D, *Background*, for more information.)

The basic concept is simple, but it proved to be several years ahead of its time. This publication explains the concept and its rationale, and calls upon industry—both the suppliers and users of information technology products—to bring it to life. Additionally, we—the current participants in the committee—outline the steps needed in the IT products industry and offer some scenarios of tag application and use.

Some of the many expected benefits of the EIT to user enterprises:

- Significant reduction in the costs of processing discovery requests;
- Significant reduction in the costs of processing Freedom of Information Act (FOIA) requests;
- Better control over trade secret information;
- Better control over privacy information;
- Better implementation of role-based security;
- Better mapping of security processes to information holdings;
- Significant reduction in the costs of course-of-business record-keeping;
- Faster access to information anywhere within the enterprise;
- More effective implementation of intellectual property policies;
- Enhancement of transaction authentication and verification;
- Better enforcement of policies for preserving information;
- Better enforcement of policies for purging information;
- More effective sharing and reuse of information.

See Annex A for illustrations of these benefits.

This report focuses on information that exists electronically because the committee believes that the greatest need and economic business case for it is with electronic information. However, the concept also can be used with non-electronic (physical) information, as discussed in Annex D.

### 3 Uniqueness of the concept

What makes the EIT proposal unique is its universality. The standard-format EIT can be used with various types of information in all forms of media, governed by any rules, for any purpose. It is a single tool for all aspects of information management. It is usable for privacy management, security management, knowledge management, intellectual property management, and evidentiary management.

As discussed below and in the Annexes, the EIT is not a panacea in a vacuum because it must be used together with management policy formulation, project planning and implementation, and regular auditing. What it offers is a single tool for a consistent information management strategy implementation throughout an enterprise. All enterprise IT personnel would need to learn this one tool to be able to implement the enterprise's information management policies in a consistent way in all corners of the enterprise.

## 4 Acronyms and definitions

### **COTS**

#### **Commercial-Off-The-Shelf**

information technology vendor products made for many different customers

### **DBMS**

#### **Data Base Management System**

computer program that organizes and provides access to a database of formatted data elements and items

### **DMS**

#### **Document Management System**

software or system used to manage the storage of and access to electronic documents

### **EDI**

#### **Electronic Data Interchange**

standardized method of exchanging predefined formatted data transactions between computers

### **EIT**

#### **Enterprise Information Tag**

electronic tag applied to/associated with an information object by a particular enterprise

### **ERP**

#### **Enterprise Resource Planning (System, Project, Product)**

software employed to integrate the information used by an organization's many different functions and departments into a unified system

### **Evidence**

documents or information presented in judicial or regulatory proceedings to support an assertion

### **FOIA**

#### **Freedom of Information Act**

federal or state statute that provides the right to request access to government agency records or information and requires government agencies to disclose requested information unless specifically exempted by the statute

### **IO**

#### **Information object**

electronic item of information or set of digital bits that has a unique, discrete existence and can be identified for evidentiary purposes

### **IT**

#### **Information Technology**

general term referring to the handling of electronic information and the resources devoted to that handling

**Metadata**

data that identifies and describes information objects

**Repository**

electronic storage facility for information

The term is used with a modifier to indicate the kind of information stored, e.g., "metadata repository."

**XML****Extensible Markup Language**

markup language used to indicate the structure, content, and rendering of data, documents, and Web pages

## 5 The concept

### 5.1 An infrastructure strategy

The EIT is not just a concept about associating metadata tags with information objects according to a standardized formula. It is a concept that unites the tags with an enterprise strategy and tactical implementation for metadata management. The tags without the metadata management support could still have some use and value, but it would be a pale shadow of the use and value when implemented together with a metadata management activity. The very concept of the tags assumes and requires the pursuit of an enterprise metadata management strategy. Because the tags are so intertwined with a parallel metadata management activity, the combination is best seen as an information management *infrastructure*.

The other dimension that makes this report an infrastructure document is that the tags can be used by software of almost every kind to make decisions about how to treat information objects. Just as tag creation can be built into off-the-shelf commercial software products, so can tag inspection or interrogation. Thus, the EIT is an infrastructure proposal in two respects: as an integrated combination of tags and metadata management, and as a uniform way to identify and describe information objects combined with software action decisions based on tag contents. Those actions could be about privacy, security, property rights, etc.

Metadata management requires awareness, identification, and description of information holdings. It is perhaps the single most important aspect of information and data architectures, or the information and data layers in IT architectures. Any enterprise that is pursuing information and data architectures probably is pursuing also a metadata management strategy to implement them. Such enterprises are already on their way to being able to take advantage of EITs because the tags will leverage the metadata management activities already underway.

Annex C speaks to some aspects of metadata management that are important to and used by EITs. Specifically, it emphasizes the databases, called repositories, that store metadata information.

### 5.2 Information objects (IOs)

An information object is any electronic item of information that an enterprise wants to manage for purposes of:

- knowing it exists;

- being able to find it;
- being able to benefit from it;
- being able to apply privacy, security and other policies to it;
- being able to preserve it; and
- being able to remove it.

An information object might be a document, an email message, a digital photograph, a sound clip, a scanned image, a spreadsheet, a voice mail message, a map, a software program, a data record, a data item, a slide presentation, or any other set of digital bits that has a unique, discrete existence and can be identified for possible evidentiary purposes. It can be stored on any electronic storage medium in which information is stored digitally. It can be on-line to a network, or off-line in a desk drawer, handbag, or briefcase. Its unique existence means that it can be stored and communicated independently of other objects and it has, therefore, some logical reason for existing by itself. Its value might be contingent upon one or more other IOs with which it is used, but its existence and its ability to be identified, stored, found, and used are not tied entirely to some other IO. Some IOs might be tied to business process software that provides semantic understanding and context, while others might derive evidentiary value by their relationships to other IOs.

IOs can be self-contained objects such as photographs, sound recordings, images, text documents, video clips, messages, etc., or they can be sets of data items stored in data base management systems (DBMSs). (Sometimes the former are called "unstructured" information, while the data items are called "structured" information.) Because data items stored in DBMSs usually are integral to business application systems and are managed carefully for evidentiary control, there might be a lesser business case for an enterprise to seek to apply EITs to them as pieces of formatted data, particularly in consideration of the careful metadata management usually associated with DBMS data stores. (The AIIM Technical Report 31 deals with the evidentiary reliability of such recorded information.) Instead of focusing on individual data items in DBMSs, there could be a strong business case to associate EITs with the transactions giving rise to the data items—for example, EDI, forms-fill, and workflow messages. Therefore, as a practical matter, the following discussion is best read in terms of self-contained objects, business transactions, and transaction messages.

Outside the DBMS-managed databases, as huge as they may be, there is a vast terrain of memos, letters, document images, email messages, voice mail messages, engineering and design drawings, maps, marketing materials, web pages, instruction and training materials, legal documents, financial reports and statements, policy manuals, multi-media training materials, videos, calendars, task lists, sound recordings, etc. All of these are or will be electronic—stored in devices or on media for which the enterprise can be held accountable—and all are candidates for tagging. All of them are IOs covered by this report.

(For additional discussion of information objects, including non-electronic items, see Annex C.)

### 5.3 Enterprise Information Tag (EIT)

The Enterprise Information Tag (EIT) is a tag applied to every information object that an enterprise chooses to tag. The enterprise will use the tags to:

- find information;
- get more value from information;