METADATA-DRIVEN INFORMATION SECURITY MODEL FOR ENTERPRISE CONTENT MANAGEMENT

A Thesis

Presented to

the Faculty of the Department of Information and Logistics Technology

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

Information Systems Security

By

Barry Yim

May 2018

METADATA-DRIVEN INFORMATION SECURITY MODEL

FOR ENTERPRISE CONTENT MANAGEMENT

Barry Yim

APPROVED:

R. Christopher Bronk, Ph.D.
Committee Chair
Assistant Professor, Information and Logistics Technology

Wm. Arthur Conklin, Ph.D.
Associate Professor, Information and Logistics Technology

Stuart Wagner, MBA
Enterprise Products Partners L.P.

George Zouridakis, Ph.D.
Associate Dean for Research and Graduate Studies, College of Technology

Daniel M. Cassler, M.A.
Department Chair, Information and Logistics Technology

ACKNOWLEDGEMENT

A decade ago today, I would have never considered writing a master thesis. Even if I tried, I would not have a clue as to where to start. Today, I am here writing an acknowledgement statement for my thesis, just couple weeks away from my graduation. It was definitely a challenging experience, and I enjoyed every moment of the journey. However, none of this would have been possible without the support that I received.

I want to start by thanking Dr. Wm. Arthur Conklin and Dr. R. Christopher Bronk. They sacrifice their nights and weekends to educating a new generation of information security professionals. No words can describe this commitment and dedication. The respect that I have is immeasurable and my appreciation is endless. Without their support and guidance, this thesis would still be a stack of blank white papers.

I want to thank Mr. Stuart Wagner, whom not only influenced me to take the leap to study information security, but also introduced me to this specific graduate program. Without his mentorship, none of this would have even happened.

I want to my family and friends for being very supportive throughout the past years. I also want to give a special thanks to a fellow classmate, Annie Jamshed, whom had spent countless hours talking me through the various challenging moments while writing this thesis.

Finally, I want to thank my wife, Nicole, for the never-ending support and for believing in me. Throughout this journey, she was always by my side, motivating me, and giving me the confidence to achieve goals and overcome challenges.

ABSTRACT

This thesis focuses on the topic of defining how a metadata-driven security model can provide an organization with the ability to dynamically manage information security for enterprise content management systems, through the use of attribute-based access controls. This thesis will cover a range of topics, such as information management, information security concepts, information taxonomy, metadata, access control, and identity management. We will review a conceptual design on the technical aspect of this model, along with studying the historical attempts and outcomes. At the end of this thesis, we will demonstrate the value of implementing a metadata-driven security model, and specifically how it can be implemented for enterprise content management.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

1 INTRODUCTION

Information is the core of our humanity. Documented information is how we know our history, how we can learn from our mistakes, remember and pass down what we have learned to share with future generations to come. Information will out-live us all.

The phrase “Information Security” covers many domains of knowledge and occupations. In terms of the general public, most people will translate this phrase and recognizes it as cyber security, network security, or data security. All of which is correct, as they make up part of Information Security. But how do we attempt to secure the vast sea of virtual information that is growing at an exponential rate, such as documents and records within a company. This is the area of focus that we will call “Content Security.”

We simply cannot just store information without finding a way to securely do so. Before computers, we had libraries with restricted sections through physical access controls, and stone tablets hidden in secret tunnels that only a few can find. Now, we have virtual information that is stored in computers throughout the world, thus securing these information is much more complex. Protecting information is just as important as being able to share information, therefore designing a dynamic security access model for content management is an important step to the future of secured information management.

In this thesis, the focus is within the domain of managing security access controls for content management systems. Throughout this research, we will revolve around one single question.

> *Is it possible to leverage the use of metadata in Enterprise Content Management (ECM) systems to dynamically manage access control, without compromising confidentiality, and maintaining the integrity of official business records?*

1.1 THESIS OBJECTIVE

The objective of this thesis is to demonstrate the concept to design and implement a content security model. Metadata-driven dynamic security rules will be used to manage access control for unstructured content, such as Word or PDF files, using metadata attributes stored in content management and identity access management systems. The goal of this thesis is to provide a fundamental understanding of metadata, taxonomy, enterprise content management systems, identity management, and access control. Along with that, we will research this topic and determine if such framework or technology currently exist. By integrating an ECM system with a metadata-driven security model, the outcome is the ability to create access control lists that can dynamically manage user permissions to unstructured business content.

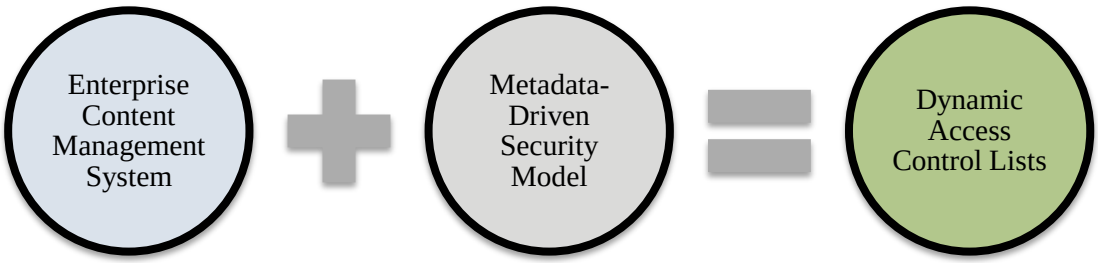

Figure 1-1 Equation to Dynamic Access Control

This thesis will include an assessment of business values to implementing this model, and demonstrate its benefits for an enterprise organization. By designing and implementing a sustainable metadata-driven security model, we can potentially reduce the cost and risk for an enterprise organization in regard to managing content security. Any industry should be able to use this as a framework to implement a dynamically secured content management system, meet business requirements, and comply with various regulatory standards.

1.2 CONCEPT OF METADATA-DRIVEN SECURITY MODEL

The concept of metadata-driven security model for any enterprise content management (ECM) systems is to allow the metadata to do the work and dynamically manage access control. Instead of reinventing the wheel, a metadata-driven information security model for content management should leverage an access control model known as “Attribute-Based Access Control” (ABAC). In chapter 2, we will further explain the ABAC model with more technical details. By using metadata attributes assigned to contents and users, dynamic security rule can be implemented to systematically maintain user access without manually updating security groups. As such, the metadata attributes that are being used for designing these security rules must be standardized and centrally managed as a single source of data. Figure 3-1, as shown below, illustrates the metadata-driven security model at a high level overview.

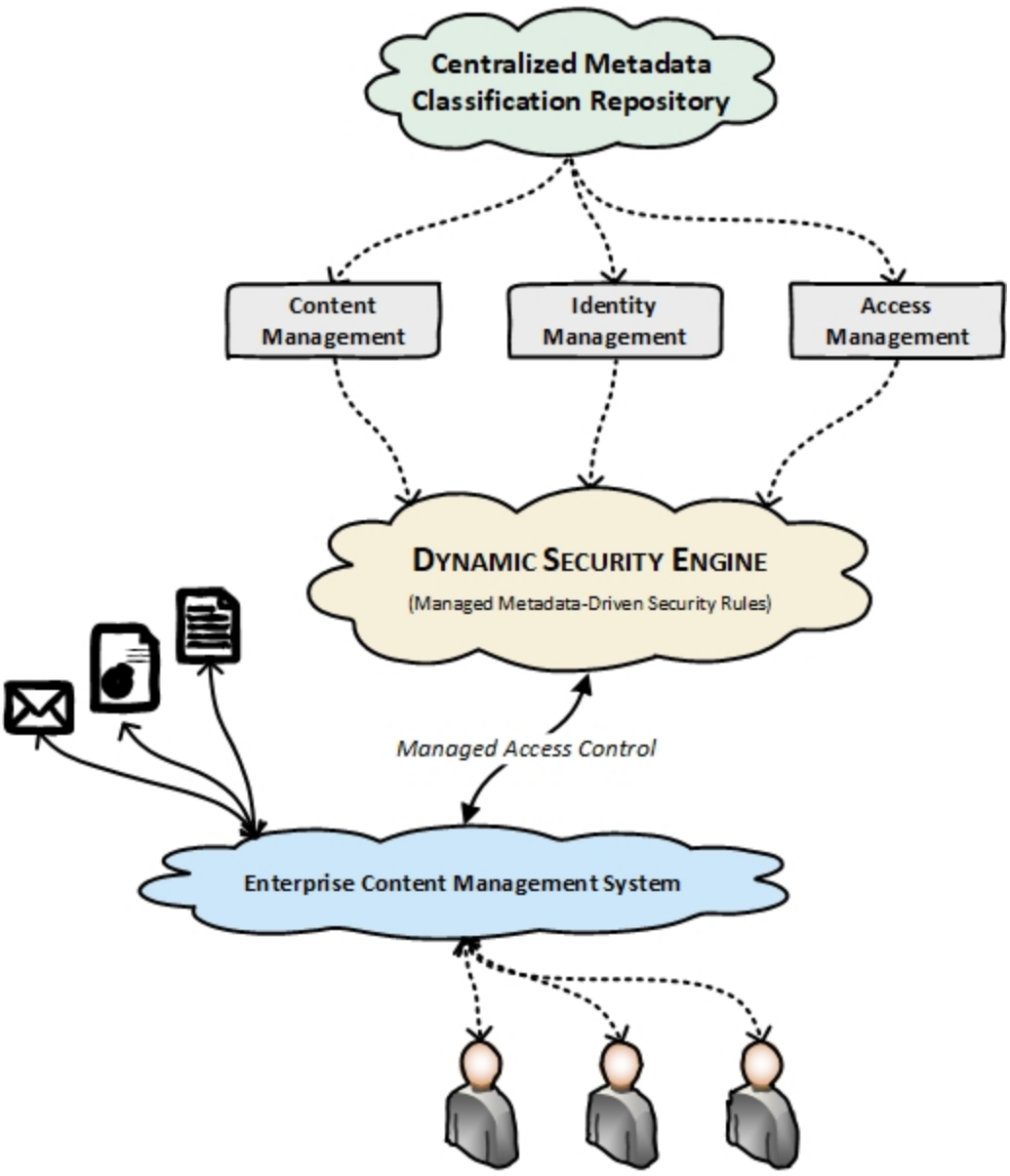

Figure 1-2 Metadata-Driven Security Model Overview

As show in the figure above, a centralized metadata classification repository provides managed metadata values for a company or organization to use across different processes, such as managing access control, managing user identity, and managing business content. With standardized metadata being used across these processes, it enables us the option to write dynamic security rules based on the metadata attributes assigned to content, users, or access

groups in a “dynamic security engine.” Thereafter, any ECM systems can then authenticate users and authorize access based on the security rules defined in the dynamic security engine. In chapter 2, we will have a literature review on each of these core components to establish a baseline of knowledge before further explaining the model in chapter 3.

From a business perspective, this conceptual metadata-driven security model may seem fairly simple, but the actual implementation can be quite complex. This security model will combine Enterprise Content Management (ECM) with Access and Identity Management (AIM), and bound them with a well-defined Information Management Program.

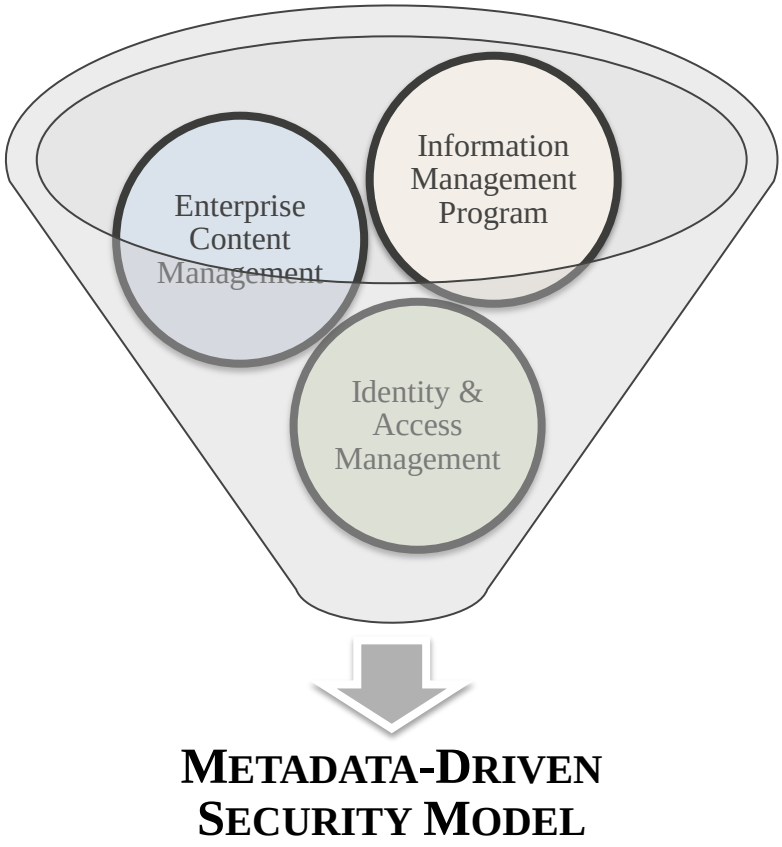

Figure 1-3 Components of Metadata-Driven Security Model

ECM has a primary responsibility to manage and store content, such as business records, and allow users to easily discover and retrieve them. It also has a responsibility to manage the content lifecycle from creation to disposition. Therefore, a successful ECM implementation will typically require an in-depth understanding of the entire business model, along with information lifecycle for every organization and information asset within the company. The Information Management Program, in this case, is responsible for covering the business aspect and creates a bridge between the business requirements and the technical requirements for ECM. It is also responsible for the governance and support of the centralized metadata classification repository, in which provides metadata classifications for both ECM and IAM. IAM provides the ability to use metadata to classify user identities and define access control rules, which will become the supporting beams to enforce confidentiality and integrity. By using the same standard metadata classifications as ECM, provided by the information management program, the consistency of metadata attributes between ECM and IAM will allow us to create metadata-driven security rule that is applicable for ECM to dynamically manage its content security.

Finally, this metadata-driven security model should be generic enough to be applied to any industries. In this thesis, we will be using business scenarios from the energy industry as our examples. Our example will focus on the needs of Commercial organizations within a company to store and retrieve commercial contracts, as well as share information or restrict access with other business units within the company. We will demonstrate how metadata-driven security rules can meet these business requirements dynamically through the use of metadata classifications.

1.3 THESIS OVERVIEW

This thesis is organized as shown below.

Chapter 1:

This chapter will provide you with an overview of this thesis and the motivation for choosing this topic.

Chapter 2:

This chapter is a literature review to provide published information in regard to attribute based access control (ABAC). In addition, this section provides background information on topic around information management, which is crucial, as this thesis is focus on the application of metadata-driven security for content management.

Chapter 3:

This chapter is the core of this thesis. It will walk us through the metadata taxonomy design to create and define a series of sample datasets. It will also use these sample datasets to conceptually design dynamic security rules for various business scenarios.

Chapter 4:

This chapter will provide you with some of the historical studies in regard to attribute based access control model, its value, and what it takes for a business to adopt it for managing content.

Chapter 5:

This chapter is the conclusion, in which will summarize our thesis research and overall finding. After reading chapter 1-4, you will be able to understand this concept and come to the conclusion that a dynamic security rule for managing user access to content is the path to the future.

CHAPTER 2

2 LITERATURE REVIEW

In this literature review, we will cover a variety of different subjects that will ultimately work together to support a metadata-driven security model. First and foremost, we will have an overview of information security concepts as it relates to a metadata-driven security model. This will define and explain terminologies, such as access control, authorization, and authentication. As this thesis is focused on using metadata to properly secure access to business content and managing documents and files, we will also include some background discussions around information management. This includes enterprise content management, electronic records management, taxonomy and metadata management, and identity and access management. As we review these topics, we will begin to realize how each component will have a role in implementing a metadata-driven security model for managing enterprise content.

2.1 INFORMATION SECURITY OVERVIEW

Information security for managing content revolves around two main processes: authorization and authentication. In this section, we will define these two terms and establish a foundation of knowledge for this thesis. Authorization is the process of permitting or denying access to resources. (Conklin, White, Williams, Davis, & Cothren, 2016) In this thesis, resources will refer to the content being managed or stored inside a centralized repository. Authentication

is the process of validating the identity of the users who are trying to gain access to a specific content. (Conklin, White, Williams, Davis, & Cothren, 2016) Authorization involves the process of defining access controls and providing a system with the ability to use access control lists or groups to properly grant or deny access. Access control is "the ability to control whether a subject can interact with an object. Subject refers to an individual or process, while object refers to a file or hardware." (Conklin, White, Williams, Davis, & Cothren, 2016) In this thesis, subject will be the employees of a company, and object will refer to any unstructured content such as documents, files, etc.

Access control is a mechanism to provide the ability of authorization to content. In order to grant access to users, we usually create groups or users which are called, security groups. Security group is a collection of subjects in which will be granted or denied access based on the security rules in place. For instance, when it is a document, we typically think of it as two security groups: Read or Read/Write. A Read group, also known as Read-Only group, means that a group of users are only granted read only permission to a file, while a Read/Write group will provide the users the permissions to edit the file. Security groups can be defined in various ways as well. Some security groups are statically managed by manually assigning users to the group one-by-one. Other security groups are dynamically managed, and are achieved by using metadata attributes associated to an employee to create dynamic security groups.

Access control can be implemented in various ways by following different models and processes. The following is an overview of the models, but we will not be diving into details of most of them in this thesis. There are confidential models such as Bell-LaPadula or Brewer-Nash security models. (Conklin, White, Williams, Davis, & Cothren, 2016) These models are designed to focus on protecting the confidentiality of an object. There are also integrity models such as

Biba security model. (Conklin, White, Williams, Davis, & Cothren, 2016) Integrity models focus on maintaining integrity of the content. Finally, there are Mandatory Access Control (MAC), Discretionary Access Control (DAC), Role-Based Access Control (RBAC), Rule-Based Access Control, and Attribute-Based Access control (ABAC). (Conklin, White, Williams, Davis, & Cothren, 2016) In this thesis, we will be focusing on the use of ABAC in managing unstructured content.

Attribute-based access control (ABAC) is a "new access control schema based on the use of attributes associated with an identity. These can use any type of attributes (user attributes, resource attributes, environment attributes, and so on), such as location, time, activity being requested, and user credentials." (Conklin, White, Williams, Davis, & Cothren, 2016) This is usually implemented under a discipline known as "Identity and Access Management". Identity and access management is "the security discipline that enables the right individuals to access the right resources at the right times for the right reasons. IAM addresses the mission-critical need to ensure appropriate access to resources across increasingly heterogeneous technology environments, and to meet increasingly rigorous compliance requirements. This security practice is a crucial undertaking for any enterprise. It is increasingly business-aligned, and it requires business skills, not just technical expertise." (Gartner, 2018) "Enterprises that develop mature IAM capabilities can reduce their identity management costs and, more importantly, become significantly more agile in supporting new business initiatives." (Gartner, 2018)

In additional to using ABAC for authentication, we will discuss the use of metadata associated to unstructured content and how these attributes will, together, play a role in creating dynamic security rules to support this metadata-driven security model for content management.

2.2 INFORMATION MANAGEMENT

"Information, as we know it today, includes both electronic and physical information. The organizational structure must be capable of managing this information throughout the information lifecycle regardless of source or format (data, paper documents, electronic documents, audio, video, etc.) for delivery through multiple channels that may include cell phones and web interfaces." (AIIM, 2008) The overall arching information management program typically includes enterprise content management, eletronic records management, identity and access management, and information taxonomy and metadata management. Each one of these core elements are essentical to creating a metadata-driven security model for cotnent management.

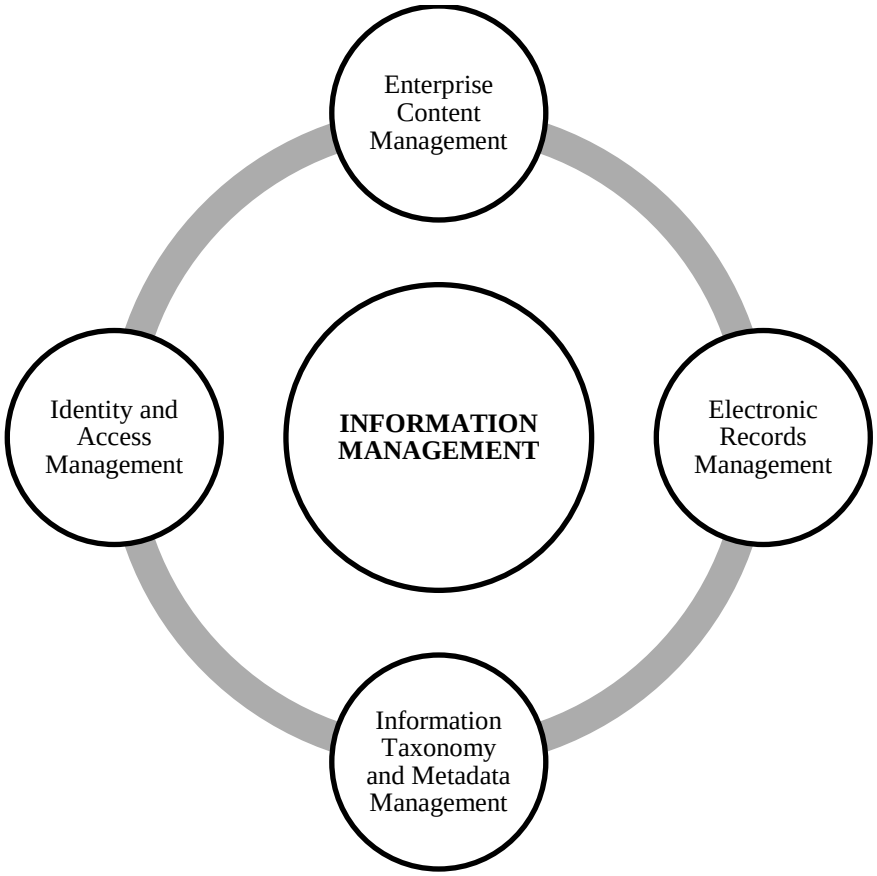

Figure 2-1 Information Management Overview

2.2.1 HISTORY OF INFORMATION MANAGEMENT

Whether it is as early as 3,000 BCE, with the first temple rooms invented by the Sumerians to store clay tablets with cuneiform scripts (Khan Academy, 2018), or looking into the modern day's Library of Congress in Washington D.C., mankind has always had the need to create, sort, file, store, archive, and index information for generations to come. In the 1980s, computers slowly came to revolutionize how we process and manage information.

"In March 1989, Tim Berners-Lee submitted a proposal for an information management system to his boss, Mike Sendall. 'Vague, but exciting', were the words that Sendall wrote on the proposal, allowing Berners-Lee to continue." (CERN, 2018) His proposal "concern[ed] the management of general information about accelerators and experiments at CERN. It discusses the problems of loss of information about complex evolving systems and derives a solution based on a distributed hypertext system." (Berners-Lee, 1989) This proposal by Tim Berners-Lee ultimately led to the creation of the World Wide Web. In essence, if the original concept for Information Management was never proposed, then the need for "cyber security" today may not exist.

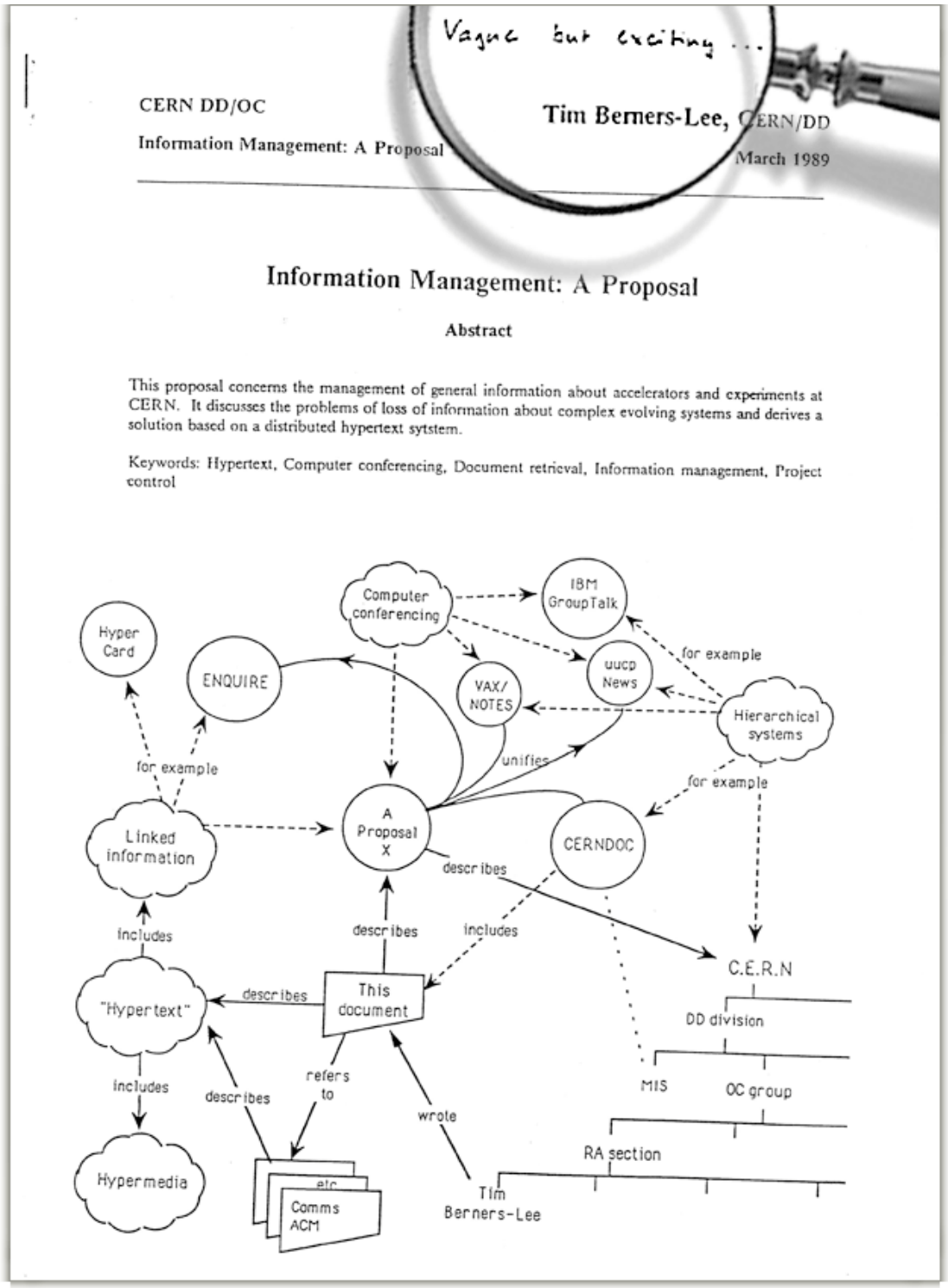

Figure 2-2 Tim Berners-Lee's Proposal (CERN, 2018)

2.2.2 BUSINESS NEEDS FOR CENTRALIZED INFORMATION MANAGEMENT

Computers were designed to compute mathematical equations, process transactions, automate or simplify processes, and help improve efficiency in everyday life. Computers are also used to create documents and records for personal or business needs. Thus, two types of information are kept in the technology world, “Structured” and “Unstructured” data. Structured data are typically refers to data stored in database, while unstructured data refers to folders and files. Whether you are using Microsoft Windows, Linux, or Mac Operating systems, at the end of the day, data and files are both created and stored. These data and output generated from information systems and processes are often important and valuable to us. We classify them as “Information Asset.”

As technology advanced over the last three decades, both processing power and storage grew exponentially. With the amount of files and content increasing at an exponential rate, managing folders and files became chaotic. In an enterprise environment, millions of folders and files are being created in silos and managed without standards. This can be problematic for an organization within company to collaborate or share information with one another. As these problems ballooned over the last decade, the demand for centralized information management also increased.

This ultimately led to phrase “Enterprise Content Management” in which was used to describe the process of centralizing information assets into a single repository in an Enterprise environment.

2.3 ENTERPRISE CONTENT MANAGEMENT

In the technology world, the term “Enterprise Content Management Systems” is frequency misinterpreted and mixed up with the term “Content Management Systems.” However, these two have very different meanings. The term “Content Management Systems” (CMS) is much more comprehensive and broader in scope. (Comentum, 2010) CMS is used to describe any systems that are used to create, edit, and manage content, in which “content” represents any unstructured data. In one perspective, we can classify an Enterprise Content Management System as a subset of CMS. Although, in recent years, the term CMS has become more adopted to specifically describe Web Content Management Systems, such as Drupal, Joomla, Word Press, etc. (Comentum, 2010)

The term “Enterprise Content Management” (ECM) is more specific and can be defined as a program that consists of processes, procedures, and technologies used in conjunction to manage unstructured content. (Comentum, 2010) In this case, the unstructured content often refers to any business documents or records. Thus, an ECM program is typically highly intertwined with an Electronic Records Management (ERM) program, which we will discuss later. ECM systems focus on storing enterprise content, such as business records, in single centralized repository for easily search and retrieval. It allows the use of metadata to support dynamic search and retrieval. We will discuss metadata in the next section.

> *“Enterprise Content Management is the systematic collection and organization of information that is to be used by a designated audience – business executives, customers, etc. Neither a single technology nor a methodology nor a process, it is a dynamic combination of strategies, methods, and tools used to capture, manage, store, preserve, and*

> *deliver information supporting key organizational processes through its entire lifecycle." (AIIM, 2018)*

2.3.1 THE EVOLUTION TO ECM

In the 1990s new generations of systems were introduced to manage digital documents created on Microsoft Office and other application packages. These systems were designed to manage active, changing documents and were referred to as electronic document management (EDM) systems. They managed all the documents in a repository or library and assigned index data or metadata to each document to manage the relationships between them and provide access control. (CIMTECH, 2007)

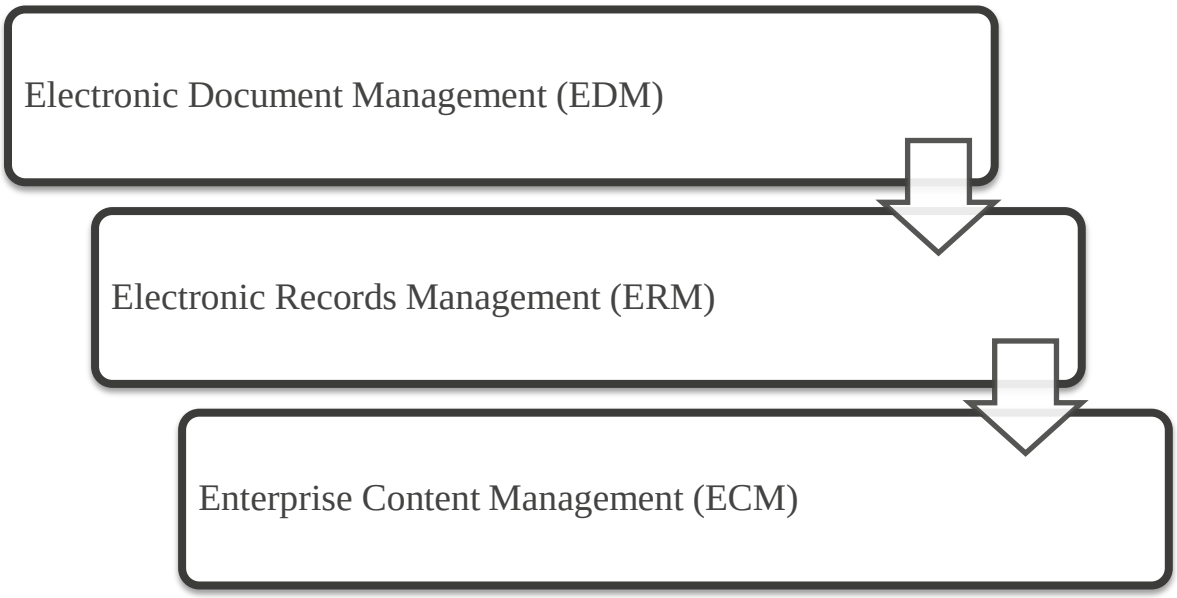

Figure 2-3 The Evolution to ECM

After EDM systems were introduced, software companies began to integrate EDM systems with ERM. ERM systems provided the abilities to classify content and leverage metadata to help manage official business records through an information lifecycle. ERM

systems provide the ability for managing records retention and disposition. EDM and ERM together, ultimately transitioned the market into the new electronic document and records management (EDRM) systems. EDRM focuses on identification, preservation, and discovery of documents. Implementing EDRM systems with an Information Management Program along with data governance and security controls marks the beginning of ECM.

2.3.2 ECM SYSTEMS OVERVIEW

In an ECM system, users typically access it through a web portal. The system is designed for users to upload documents and files into a centralized file server. The file server is connected to an ECM database, in which it will store a series of metadata attributes for each file. These metadata elements can be generated systematically by the ECM system or manually assigned by the user during the check-in or upload process. During the upload process, the user has ability to fill out a metadata form, which is typically a pre-defined metadata form with fields applicable to each content type. Some metadata fields use a list of values managed by the company, and we call these metadata "Managed Metadata." Managed metadata catalogues, as the source data, are usually kept in a separate database to maintain data integrity and enforce change management processes. As metadata attributes are assigned to content, these values become the identity of the file or content. This is important to note for this thesis as the content is the object in which we are trying to secure. Since ECM provides the ability to classify these objects with metadata attributes and provide them with an identity, we can then leverage these values to create our dynamic security rules. (PaperFree Corporation, 2018) The figure below demonstrates an example of how information is captured for ECM.

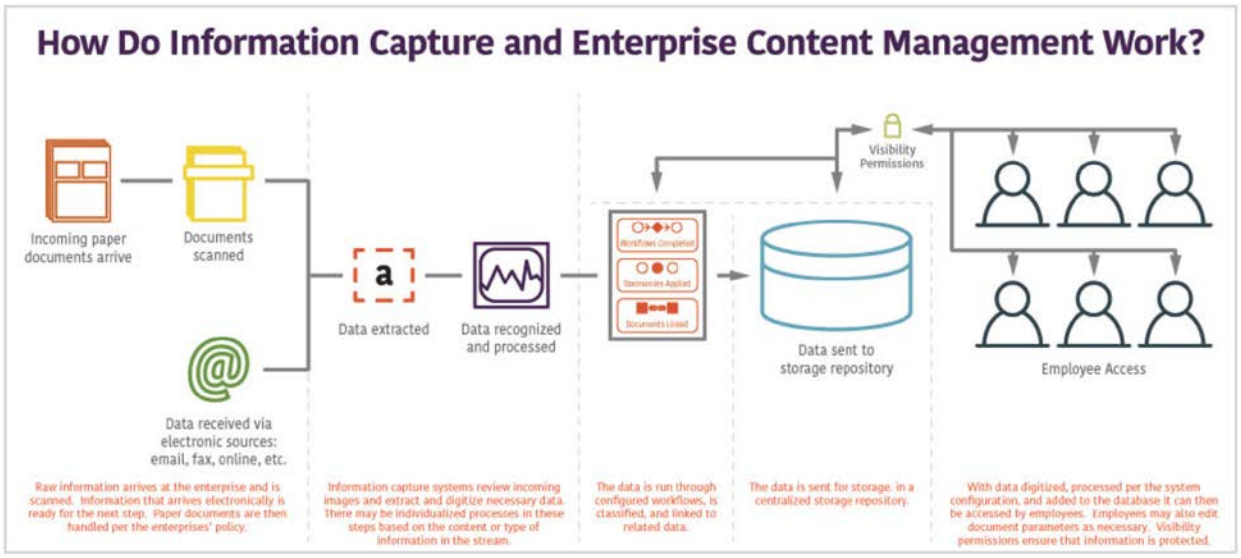

Figure 2-4 Information Capture Process for ECM (PaperFree Corporation, 2018)

2.4 INFORMATION TAXONOMY

We developed an understanding of what an ECM system is in the previous section and how it functions in an organization. Now we will cover the topic of information taxonomy to help define a centralized metadata framework that we can use for ECM. After we understand information taxonomy and ECM thoroughly, then we will discuss the details of the concept of the metadata driven security model.

Before we define information taxonomy, let's ask ourselves: What is taxonomy? When we use the term "taxonomy", we typically think biological hierarchy. Generically speaking, taxonomy is a scheme of classifications. Taxonomy provides us with a standardized and uniform way to communicate and establish consistency, by means of grouping and separating things based on its similarities and differences. They are designed based on a well-defined hierarchical framework or structure, in which the groups must be mutually exclusive. The figure below illustrates the characteristics that any taxonomy must have.

Must have these characteristics . . .

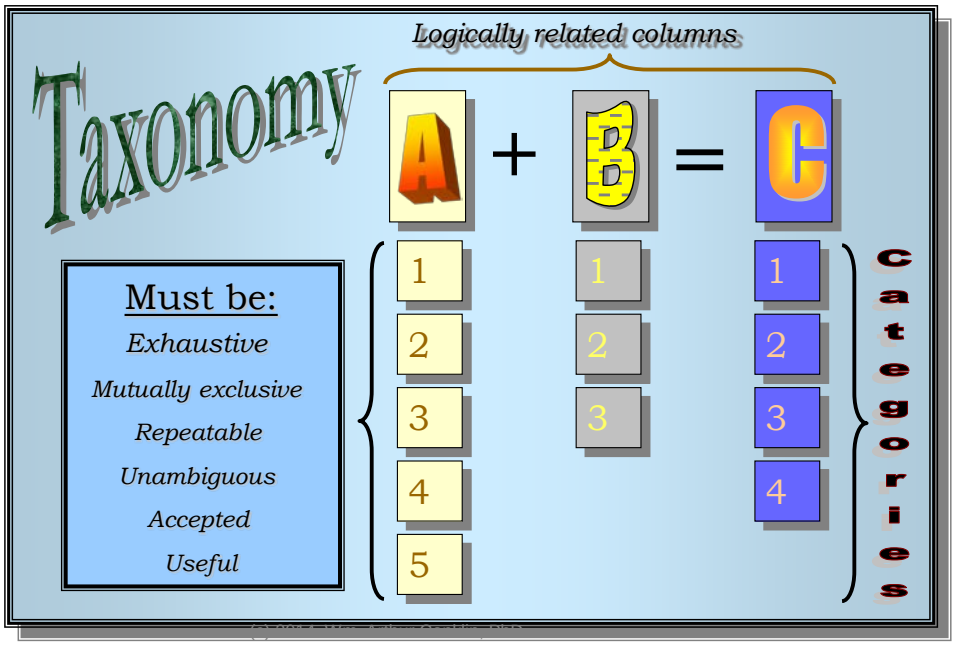

Figure 2-5 Taxonomy Characteristics (Conklin, 2014)

The above figure shows the very basic format of Taxonomy, and can be applied to any taxonomy breakdown. From this figure we can see that the start of any taxonomy classification process is to breakdown information by looking at similarities and differences. First we would define the top most parent level of logically related columns. In the example above, the logical related columns are "A, B, and C." These three would be the parent categories, under which will be more groupings of different categories. Keep in mind that the parent categories are the key similarity of each of the child categories. Focusing on the fact that each of the groups and the subgroups of similarities or differences must be exhaustive, repeatable, unambiguous, accepted, and useful, as such that, each organization must be able to come up with an appropriate information taxonomy that is suitable for their needs and requirements. The keywords used for the categories are metadata. Each group of metadata is mutually exclusive, such as content

function versus business function. Within a class of metadata, such as content function, the subset of classification are similar to its parent group.

Without taxonomy, we cannot develop common reporting criteria, develop processes, or create standardization. Taxonomy provides us with a "Common Language" to share information. (Conklin, Introduction to Incident Response, 2014) Information taxonomy is simply the science of classification of information. For information security systems, we can define a variety of taxonomy depending on the need and objective. Incident response, for example, needs to have intrusion taxonomy in order for an incident response team to communicate in the "same language." Below is a simple example of information taxonomy for Intrusions. We can clearly see the logical related columns, parent categories and child categories. We can report an intrusion incident and correctly communicate it consistently, such as (an attacker) Hackers → (used) Information Exchange → (to) Configuration → Flood → Data → (caused a) Theft of Resources → (resulted in) Damage.

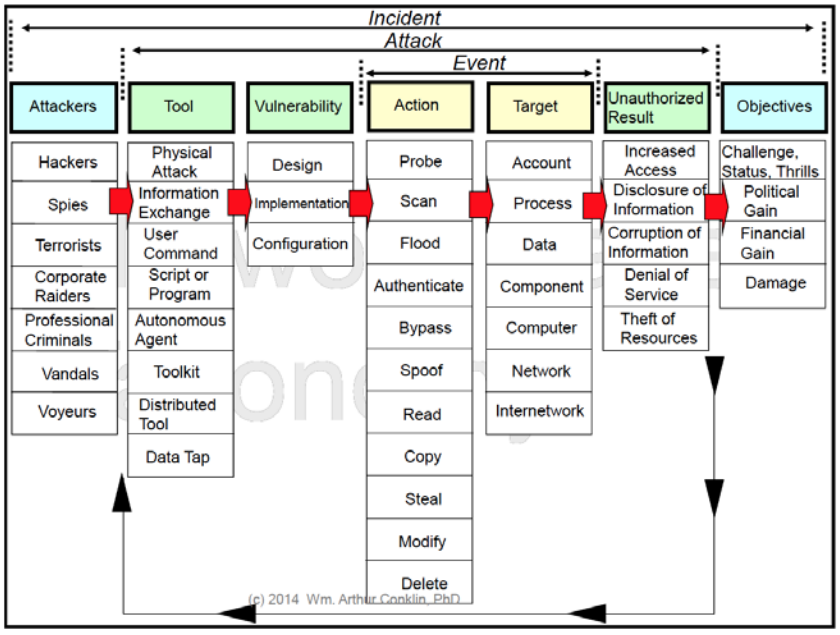

Figure 2-6 Intrusion Taxonomy (Conklin, 2014)

The figure below is an example of ECM taxonomy with multiple sub-taxonomies. The figure illustrates how ECM metadata taxonomy is designed to classify content. Content category and content type are both attributes designed for classifying content, but content type describes the content at a more granular level than content category. In theory these information taxonomy are application to any organization in any industry or sector. This information taxonomy for ECM will allow dynamic security rules to process properly, as metadata definitions are standardized across a company. A content function could be a Legal (content function) Contract (content category) → for Commercial (organization) NGL Management (business unit) → that is applicable to Company I-A (business entities, internal) and Company E-X (business entities, external). This example can be cross-referenced in section 3.2.

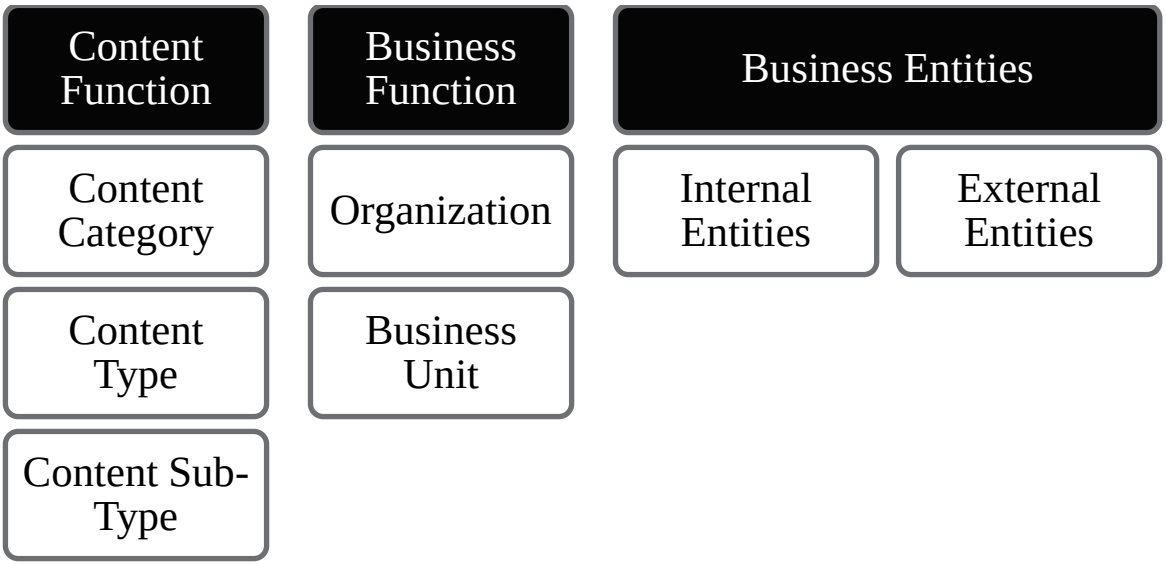

Figure 2-7 ECM Metadata Taxonomy

Information taxonomy is important in this metadata-driven security model, because it sets the foundation to allow the dynamic security rules to function properly. The taxonomy will provide us with standardization and help enforce consistency in terms of metadata utilization. Therefore, grouping information or content by means of classification will later be transformed

into a metadata framework. Populating values into taxonomies is to define the metadata of a set of information. By having a defined taxonomy and a defined set of metadata, it will make the information classifications reliable. Having a defined taxonomy also will create standardization and provide a common language for both people and machines to communicate. This is important to this thesis, as the metadata-driven security model's integrity is dependent upon the integrity of the metadata taxonomy itself. We will discuss metadata in detail in the upcoming sections.

2.5 METADATA OVERVIEW

In order to insure the integrity of information taxonomy we must understand Metadata. This section will first define the term metadata, provide history and background information, and then explain the relationship between metadata and ECM. Understanding the use of metadata in ECM will provide the knowledge necessary to defining dynamic security rules in Chapter 3.

In a short definition, metadata is simply data about data. In the computer world, metadata exists everywhere. In the case of information assets such as a document, metadata is used to help identify and describe the document.

> *"Metadata is information stored within a document that is not evident by just looking at the file. It is an electronic 'fingerprint' that automatically adds identifying characteristics, such as the creator or author of the file, the name of individuals who have accessed or edited the file, the location from which the file was accessed, and the amount of time spent editing the file. In addition to data that is automatically added to a document, there is user-introduced metadata, such as*

> *tracked changes, versions, hidden text and embedded objects." (Harvard Law School, 2018)*

2.5.1 THE ORIGIN OF METADATA

The use of metadata has a long history, and the idea did not originate in the technology era. In fact, the first documented use of metadata dates back to 280 B.C. in the Great Library of Alexandria. "Library staff under Zenodotus attached a small dangling tag to the end of each scroll, which contained information on each work's author, title, and subject so that materials could be easily returned to the area in which they had been classified, but also so that library users did not have to unroll each scroll in order to see what it contained. Obvious and unimpressive though it may seem to those accustomed to modern libraries, this was the first recorded use of metadata, a landmark in library history." (Philips, 2010) With metadata in libraries to index books and artifacts, locating and discovering of these assets became much easier and more efficient.

After the World Wide Web was introduced in the 1990s and the significant increased use of information systems increased in the 2000s. Electronic document management / enterprise content management systems were developed to help with records management by using metadata in a similar manner to its predecessor, libraries.

2.5.2 STANDARD METADATA TYPES

"The metadata is what is used to identify and retrieve the content, to organize it, control access to it and assure its integrity. If you are coming to this from the paper filing world then

think of metadata as indexing data or finding aid data plus additional data used to control and manage the content." (CIMTECH, 2007)

There are various types of metadata. Some metadata are system-defined; others are user-defined. System-defined metadata are mostly used for interoperability and preservation. Most of them are designed as hidden data elements. User-defined metadata are more commonly used for discovery and navigation. They are typically attributes that are more descriptive and used to help define content in non-technical or business terminology.

Below are examples of metadata types and properties defined by the National Standards Organization (NISO). (Riley, 2017)

Metadata Type	Metadata Type Description
Descriptive metadata	For finding or understanding a resource
Administrative metadata · Technical metadata · Preservation metadata · Rights metadata	· For decoding and rendering files · Long-term management of files · Intellectual property rights attached to content
Structural metadata	Relationships of parts of resources to one another
Markup languages	Integrates metadata and flags for other structural or semantic features within content

Table 2-1 Metadata Type Published by NISO (Riley, 2017)

Metadata Type	Example Properties	Primary Uses
Descriptive metadata	· Title · Author · Subject · Genre · Publication date	· Discovery Display · Interoperability
Technical metadata	· File type · File size · Creation date/time · Compression scheme	· Interoperability · Digital object management · Preservation
Preservation metadata	· Checksum · Preservation event	· Interoperability · Digital object management · Preservation

Rights metadata	· Copyright status · License terms · Rights holder	· Interoperability · Digital object management
Structural metadata	· Sequence · Place in hierarchy	· Navigation
Markup languages	· Paragraph · Heading · List · Name · Date	· Navigation · Interoperability

Table 2-2 Metadata Properties and Primary Uses Published by NISO (Riley, 2017)

2.5.3 METADATA CLASSIFICATION FOR ECM

The following are five groups of metadata classifications that we will focus on for this research to implement metadata-driven security for ECM systems.

- Content Type Classification
- Business Function Classification
- Business Role Classification
- Information Sensitivity Classification
- Auxiliary Classification

Content Type, Business Organization, and Information Sensitivity classifications are equally important, as they are mandatory for designing the dynamic security rules. Auxiliary classifications, on the other hand, are typically optional values used as descriptive metadata for to ease the discovery of information. However, depending on the requirements, they can also be used for creating the dynamic rules.

In reference to the “Standard Metadata Types” published by NISO, as discussed in Section 2.5.2, the table shown below defines the metadata type of each of the metadata classification groups.

Metadata Classification Groups	Metadata Type
Content Type Classification	· Descriptive metadata
Business Function Classification	· Descriptive metadata
Business Role Classification	· Descriptive metadata
Information Sensitivity Classification	· Descriptive metadata
Auxiliary Classification	· Descriptive metadata · Structural metadata

Table 2-3 Metadata Classification Group Types

Since our security model is designed for ECM systems to manage unstructured data, most of the metadata fields used here are descriptive metadata. The advantage of using descriptive metadata is to allow the business to manage these data independent of any vendors or software, therefore allowing the ability to position the metadata taxonomy at the level of master data management.

These metadata classification groups are also meant to be used interchangeably across the technical areas of content management, identity management, and access management. In chapter 3, we will create sample metadata and apply them to each of these technical areas to demonstrate the concept of creating metadata-driven security rules for managing content. The figure shown below illustrates how the various classification groups are applicable to each technical area.

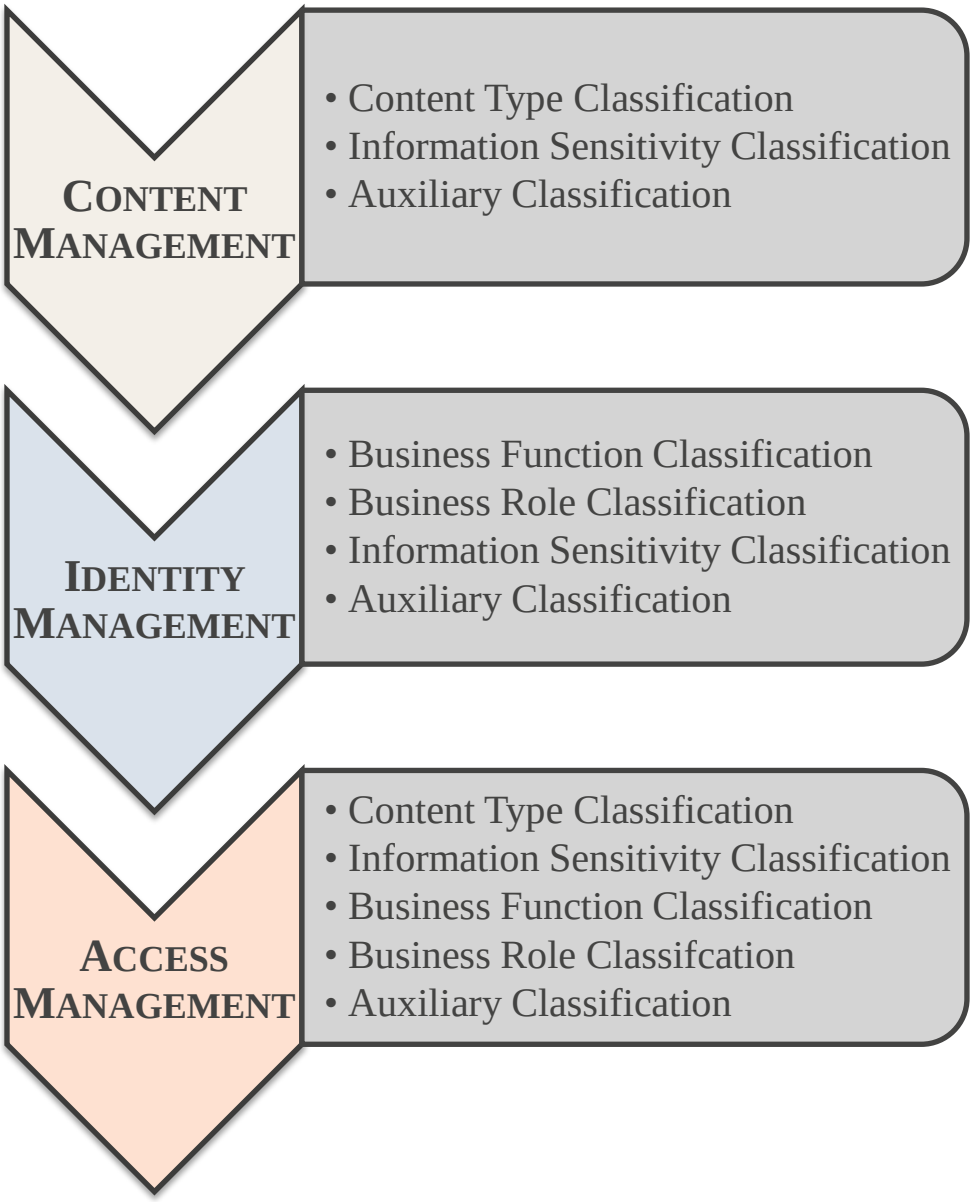

Figure 2-8 Metadata Classification for Content, Identity, and Access Management

2.5.4 GOVERNMENT USE OF METADATA FOR INFORMATION SECURITY CLASSIFICATION

Nation-states have long used metadata for information security classification. As stated by the Office of the Director of National Intelligence, “information sharing within the national

intelligence enterprise will increasingly rely on information assurance metadata (including information security markings) to allow interagency access control, automated exchanges, and appropriate protection of shared intelligence when necessary." (Office of the Director of National Intelligence, 2017) Countries and nations across the world all leverage data security classifications to protect their intelligence and national security.

Referencing the United States of America as an example, Executive Order 13526 published on December 29, 2009 by President Barack Obama, prescribes a uniform system for classifying, safeguarding, and declassifying national security information, including information relating to defense against transnational terrorism. (Obama, 2009)

> *"Our Nation's progress depends on the free flow of information both within the Government and to the American people. Nevertheless, throughout our history, the national defense has required that certain information be maintained in confidence in order to protect our citizens, our democratic institutions, our homeland security, and our interactions with foreign nations. Protecting information critical to our Nation's security and demonstrating our commitment to open Government through accurate and accountable application of classification standards and routine, secure, and effective declassification are equally important priorities." (Obama, 2009)*

In regard to records management and information sharing, Executive Order 13526 is a prime example to demonstrate how information and security classification is inseparable, when there is a need for a building a successful information security management program. The importance of having integrity in metadata is the key to the metadata-driven security model for

ECM. The next step is to understand how the records or content classifications are defined in the centralized information taxonomy based on the process to of reviewing and organizing the records repository for an organization, which is known as Electronic Records Management.

2.6 ELECTRONIC RECORDS MANAGEMENT

2.6.1 ELECTRONIC RECORDS MANAGEMENT

Electronic records management (ERM) is important, because it helps define and outline the entire record repository for an organization. In an ERM program, Record Managers from every department work together to help a company or organization create a "file plan." According to Alfresco Software, Inc., a file plan "is a container for records, folders, categories and retention schedules. It's effectively a virtual filing cabinet for storing records, and is the basic structure of Records Management. This structure lets you classify and group records with similar characteristics." (Alfresco Software Inc., 2018) Therefore, a file plan will provide the necessary information for defining Content Classification metadata. Below is an excerpt, from the Association for Information and Image Management (AIIM) organization, defining records management.

> *"ISO standard 15489: 2001 defines Records Management (RM) as the field of management responsible for the efficient and systematic control of the creation, receipt, maintenance, use and disposition of records, including the processes for capturing and maintaining evidence of and information about business activities and transactions in the form of records. Electronic Records Management (ERM) ensures*

> *your organization has the records it needs when they are needed." (AIIM, 2018)*

As you can see of this excerpt, ERM serves as the baseline to automate the manual records management process into a dynamically secured content management system. In the next chapter of this thesis, we will see how the file plan or content classifications will serve as one of the most important metadata attribute in defining security rules for ECM.

CHAPTER 3

3 METADATA-DRIVEN SECURITY MODEL

3.1 METADATA TAXONOMY DESIGN AND IMPLEMENTATION

The first step to implementing a metadata-driven security model involves designing standardized metadata taxonomy for consistent data classification sets the foundation and serves as a prerequisite for designing dynamic security rules. Without spending the time and effort to implement a standard taxonomy tailored to the organization, it can negatively impact an organization's ability to implement any enterprise content management system.

Depending on the size of the organization, the size of the metadata taxonomy framework can vary in size. A software development start-up in the health care industry compared to Fortune-100 International Corporation in the oil and gas industry will have a completely different taxonomy in all aspects. Thus, a good metadata taxonomy design must be tailored specifically to the business itself. The process involves intense requirements gathering and analysis to fully understand the information lifecycle in segment of the company.

Without going to deep into explaining how metadata taxonomies should be designed, as that is not the scope of this research, we will focus only on the metadata domains necessary for creating the dynamic security rules.

The figure shown below demonstrates how various metadata classifications can be assigned to a single content.

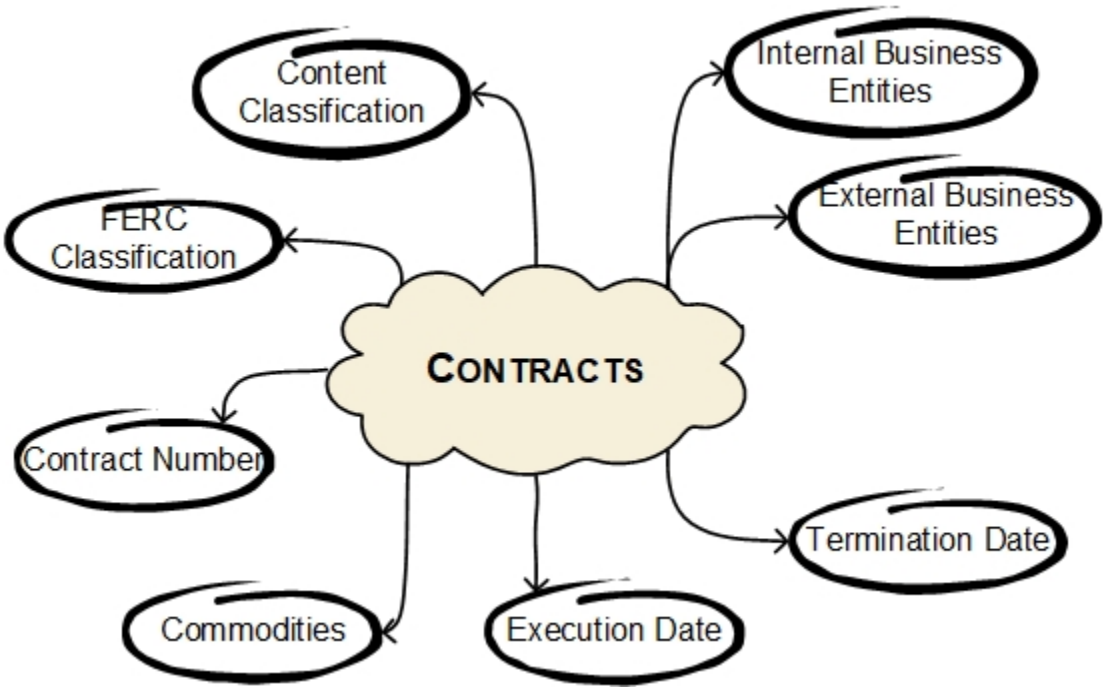

Figure 3-1 Sample Metadata Classification for Content

The figure shown below demonstrates how various metadata classifications can be assigned to a user.

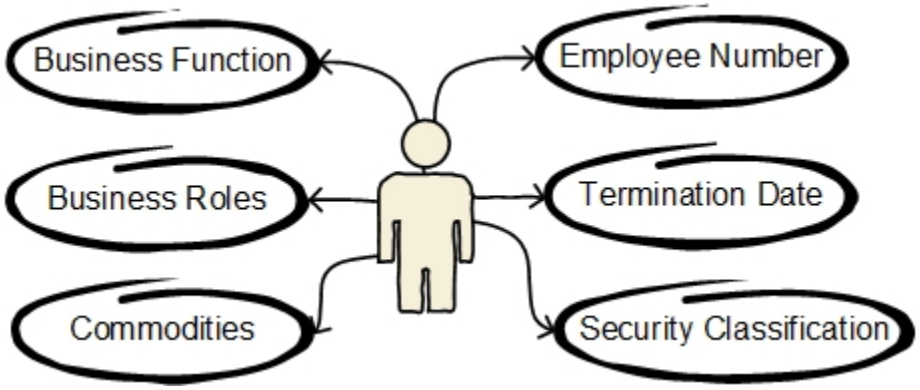

Figure 3-2 Sample Metadata Classification for User

3.1.1 CONTENT TYPE CLASSIFICATION

Content type classifications are used to identity and define the object or content. These values should be standardized for the entire organization and managed by a department commonly known as the “Records and Information Management.”

The figure shown below is an example of a taxonomy designed for content type classification.

Figure 3-3 Content Classification Taxonomy

Content Function	Content Category	Content Type	Content Sub Type
Administrative	Publications	Press Release	
Administrative	Publications	Industry Newsletters	
Environmental	Permits	Water Permit	Water Acquisition Permit
Environmental	Permits	Water Permit	Water Well Permit
Legal	Contracts & Agreements	Commercial Contract	Purchase Agreement
Legal	Contracts & Agreements	Commercial Contract	Lease Agreement
Legal	Contracts & Agreements	IT Contract	Software Licensing Agreement

Table 3-1 Content Classification Metadata Table

3.1.2 BUSINESS FUNCTION CLASSIFICATION

The business function classifications are used for identity management. The dynamic security rules will then leverage the values to group content or people by their business functions to provide departmental access control. For companies with a well-defined organization

structure, these classifications will have a one-to-one relationship. Thus, this list of values will often be managed by the Human Resource organization.

The figure shown below is an example of a taxonomy designed for business function classification.

Figure 3-4 Business Function Classification Taxonomy

Organization	Business Unit
Accounting	Payroll Accounting
Accounting	Volume Accounting
Commercial	NGL Management
Commercial	Crude Management
Information Technology	Network Infrastructure
Information Technology	Enterprise Business Systems
Legal	

Table 3-2 Business Function Classification Metadata Table

3.1.3 INFORMATION SENSITIVITY CLASSIFICATIONS

Information sensitivity classifications on average are typically broken into 3 to 4 levels, depending on the organization. Between nation-state and the private sector, the main differences in their classification are usually how they label the value.

University of Illinois at Chicago, as an example, defined 6 levels of classification. “Four classes of data are defined by the University of Illinois globally for the institution: High Risk, Sensitive, Internal, and Public. This Program additionally defines a fifth class, Sensitive Data

Collection, and an overlapping sixth class which may be present within any of the preceding five data classes, Personally Identifiable Information (PII) Data." (University of Illinois, 2018)

The following table demonstrates a sample set of data classifications based on information gathered from the University of Illinois.

Information Sensitivity Classification	Information Sensitivity Classification Description
High Risk Data (Confidential)	Information that, if disclosed or modified without authorization, would have severe adverse effect on the operations, assets, or reputation of an organization's obligations concerning information privacy.
Sensitive Data	Information that, if disclosed or modified without authorization, would have serious adverse effect on the operations, assets, or reputation of an organization's obligations concerning information privacy.
Sensitive Data Collection	A collection of Sensitive Data that results from compiling (i.e., collecting) the Sensitive Data from multiple sources.
Internal Data	Information that, if disclosed or modified without authorization, would have moderate adverse effect on the operations, assets, or reputation of an organization's obligations concerning information privacy.
Personally Identifiable Information	Personally Identifiable Information (PII) Data is any information about an individual maintained by a Unit, including (1) any information that can be used to distinguish or trace an individual‘s identity, such as name, social security number, date and place of birth, mother‘s maiden name, or biometric records; and (2) any other information that is linked or linkable to an individual, such as medical, educational, financial, and employment information.
Public	Information intended for public use that, when used as intended, would have no adverse effect on the operations, assets, or reputation of an organization's obligations concerning information privacy.

Table 3-3 Information Sensitivity Classification (University of Illinois, 2018)

The granularity of data sensitivity classifications is, once again, highly dependent on the need of an organization. In most cases, classification of Confidential, Internal Use Only, and Public is sufficient for managing content in a mid-size organization.

Information security classifications are commonly used as metadata to tag content, at a file-level or within the XML-level. In this case, the content is secured as long as the user access accounts are properly populated or maintained.

In regards to our metadata-driven security model, information sensitivity classification is just another attribute and can used to tag content or identities. By using the information sensitivity classification as metadata at both the subject (user) and object (content) level, security rules can be further refined to enforce both confidentiality and integrity.

3.1.4 BUSINESS ROLE CLASSIFICATIONS

Business role classifications are metadata attributes used to define the roles of an individual or organizations. These classifications should not be dependent on any organizational structure. They should be mutually exclusive values to objectively define a business role that can be applicable to one or many business organizations.

Business Role	Business Role Description
Contract Administrators	Creates, process, and manage contracts and agreements
Contract Consumers	Search and view contracts for application business operations

Table 3-4 Business Role Classification - Metadata Options

3.1.5 AUXILIARY CLASSIFICATIONS

Auxiliary classifications are the catch-all metadata attributes. These classifications can be defined for any subject and should be independent from all other classifications groups. In other words, auxiliary classification is a groups or collection of various metadata classifications that

can be used for search and retrieval, grouping, sorting, etc. While it is best practice to design metadata classifications at a global level, occasionally it is necessary to create content or business specific metadata. These types of classifications are all consider as part of the auxiliary classification group.

Examples of auxiliary classifications include geographical data, project data, regulatory data, asset data, etc. They may also dates and time attributes, such as execution dates, terminations dates, effective to and from dates, etc. Lastly, they can be simple as Boolean values using check boxes or “Yes/No” options.

The table shown below is a sample set of auxiliary classifications that will later be used in this research.

Metadata Field	Field Type	Managed?
Internal Business Entities	Text (List)	Yes
External Business Entities	Text (List)	Yes
FERC Classification	Boolean	Yes
Commodities	Text (List)	Yes
Contract Number	Text (Entry)	No
Execution Date	Date	No
Termination Date	Date	No

Table 3-5 Auxiliary Classification - Metadata Field Definition

Notice that there is a column titled “Managed” in the table above. This column indicates whether or not the list of values is centrally managed by the Information Governance Program. Only managed metadata fields are applicable to be used in metadata-driven security rules.

3.2 METADATA CLASSIFICATION MAPPING

In this section, we will demonstrate how these metadata classifications will be used to tag and classify various content and users.

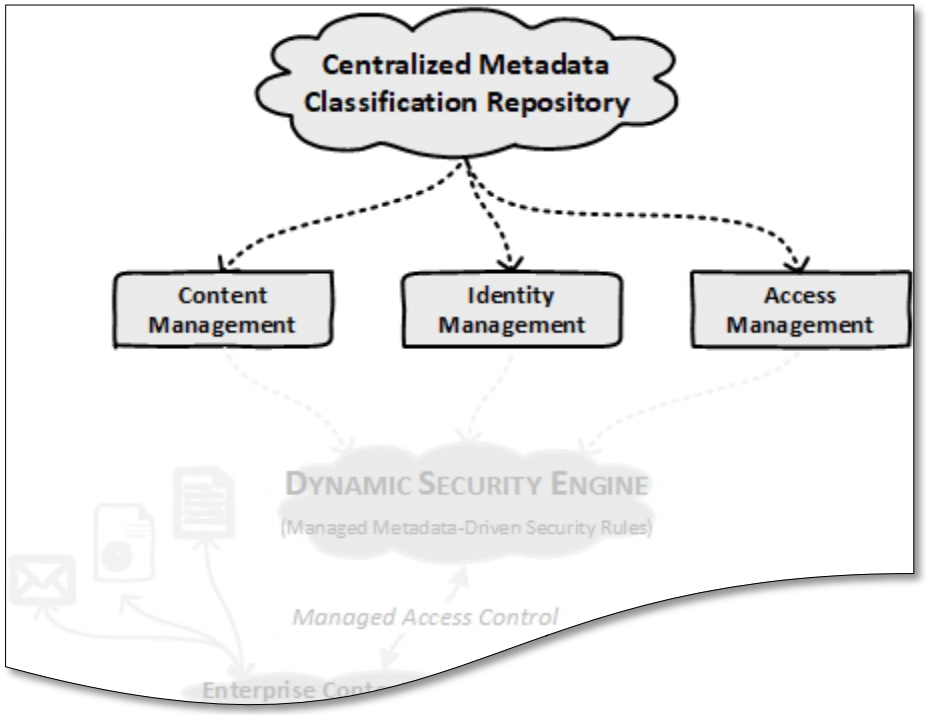

Figure 3-5 Metadata Classification Mapping Overview

3.2.1 MAPPING CONTENT WITH CONTENT CLASSIFICATION

In our business scenario to classify contracts and agreements, the content type classification taxonomy can be translated as follow:

Figure 3-6 Content Classification Taxonomy (Legal Contracts)

The table shown below demonstrates how metadata attributes for two different contracts will be stored in an ECM system.

	Sample Contract 1	Sample Contract 2	Sample Contract 3
Content Function:	Legal	Legal	Legal
Content Category:	Contracts & Agreement	Contracts & Agreement	Contracts & Agreement
Content Type:	Commercial Contract	Commercial Contract	IT Contract
Content Sub-Type:	Purchase Agreement	Lease Agreement	Software Licensing Agreement

Table 3-6 Content Classification Metadata Mapping (Contracts)

In another example to classify permits, the content type classification taxonomy can be translated as follow:

Figure 3-7 Content Classification Taxonomy (Environmental Permit)

The table shown below demonstrates how metadata attributes for two different contracts will be stored in an ECM system.

	Sample Permit 1	Sample Permit 2
Content Function:	Environmental	Environmental
Content Category:	Permits	Permits
Content Type:	Water Permit	Water Permit
Content Sub-Type:	Water Acquisition Permit	Water Well Permit

Table 3-7 Content Classification Metadata Mapping (Permits)

3.2.2 MAPPING CONTENT WITH INFORMATION SENSITIVITY CLASSIFICATION

In an ECM system, information sensitivity classifications should be configured to allow assignment at the Content Category or Content Type level. In addition, the classification at a Content Type level should take precedent and supersede the classification at the parent Content Category level.

In our example, we will use the following information sensitivity classifications.

Information Sensitivity Classification	Information Sensitivity Classification Description
Confidential	Information that, if disclosed or modified without authorization, would have **severe adverse effect** on the operations, assets, or reputation of an organization's obligations concerning information privacy.
Sensitive	Information that, if disclosed or modified without authorization, would have **serious adverse effect** on the operations, assets, or reputation of an organization's obligations concerning information privacy.
Internal Use Only	Information that, if disclosed or modified without authorization, would have **moderate adverse effect** on the operations, assets, or reputation of an organization's obligations concerning information privacy.
Public	Information intended for public use that, when used as intended, would have **no adverse effect** on the operations, assets, or reputation of an organization's obligations concerning information privacy.

Table 3-8 Information Sensitivity Classification - Metadata Options

The table shown below demonstrates how the information sensitivity metadata attributes are used to tag the Content Category.

Content Function	Content Category	Information Sensitivity Classification
Administrative	Publications	Public
Environmental	Permit	Internal Use Only
Legal	Contracts & Agreements	Sensitive

Table 3-9 Information Sensitivity Classification Metadata Mapping (Content Category)

The table shown below demonstrates information sensitivity metadata attributes are also used to tag the Content Type.

Content Function	Content Category	Content Type	Information Sensitivity Classification
Legal	Contracts & Agreements	Commercial Contract	Confidential

Table 3-10 Information Sensitivity Classification Metadata Mapping (Content Type)

Based on the rule that classification at the Content Type level takes precedent and supersedes the classification at the parent Content Category level, the metadata mapping for contracts as shown above can be interpreted as the following statement.

- All “Contracts & Agreements,” with the exception of “Commercial Contracts” are classified as “Sensitive.”
- “Commercial Contracts” are classified as “Confidential.”

3.2.3 MAPPING CONTENT WITH AUXILIARY CLASSIFICATION

In our business scenario to classify contracts and agreements with auxiliary metadata fields, we will demonstrate the use of both manual-input and auto-generated metadata fields.

Manual-input metadata fields are metadata values assigned to content manually by the user at during the check-in process. Auto-generated metadata fields are metadata values that is assigned systematically though a pre-defined metadata rule or logic.

3.2.3.1 MANUAL-INPUT AUXILIARY METADATA

Below is a list of manual-input metadata fields.

- Internal Business Entities
- External Business Entities
- Commodities
- Execution Date
- Termination Date

The table shown below is a sample list of metadata values for the “Internal Business Entities” field. Notice that there are additional attributes associated to these values. These values will be discussed in the next section.

Internal Business Entities	FERC Regulated?	Marketing?
Company I-A	Yes	No
Company I-B	No	No
Company I-C	No	Yes

Table 3-11 Auxiliary Classification - Internal Business Entities

The table shown below is a sample list of metadata values for the “External Business Entities” field.

External Business Entities
Company E-X
Company E-Y
Company E-Z

Table 3-12 Auxiliary Classification - External Business Entities

The table shown below is a sample list of metadata values for the “Commodities” field.

Commodities
Crude
Natural Gas
NGL

Table 3-13 Auxiliary Classification - Commodities

The table shown below is an example of how these auxiliary classifications are added on to the existing content classifications for the same two different contracts as shown in Section 3.2.1.

	Sample Contract 1	Sample Contract 2
Content Function:	Legal	Legal
Content Category:	Contracts & Agreement	Contracts & Agreement
Content Type:	Commercial Contract	Commercial Contract
Content Sub-Type:	Purchase Agreement	Lease Agreement
Internal Business Entities	Company I-A	Company I-B
External Business Entities	Company E-X	Company E-X
Commodities	Natural Gas	NGL
Contract Number	CCPC-NG-00165	CCLA-NGL-00335
Execution Date	11/1/1999	5/15/2013
Termination Date	7/31/2020	12/31/2035

Table 3-14 Auxiliary Classification Metadata Mapping – Manual-Input (Contracts)

3.2.3.2 AUTO-GENERATED AUXILIARY METADATA

To demonstrate how auto-generated metadata fields are populated systematically though a pre-defined metadata rule or logic, we will use a Federal Energy Regulatory Commission (FERC) requirement for restricting sensitive information as an example to auto populated a "FERC Classification" field. In this example, we will assume that we have implemented an auto-generated metadata classification rule with the following conditions to meet FERC regulation's requirements.

- If the Internal Entities field on a commercial contract does not include any "Regulated" business entities, then the FERC Classification field will be populated with "Un-Restricted"
- If the Internal Entities field on a commercial contract includes "Regulated" business entities and "Marketing" business entities, then the FERC Classification field will be populated with "Un-Restricted"
- If the Internal Entities field on a commercial contract includes only "Regulated" business entities, then the FERC Classification field will be populated with "Restricted"

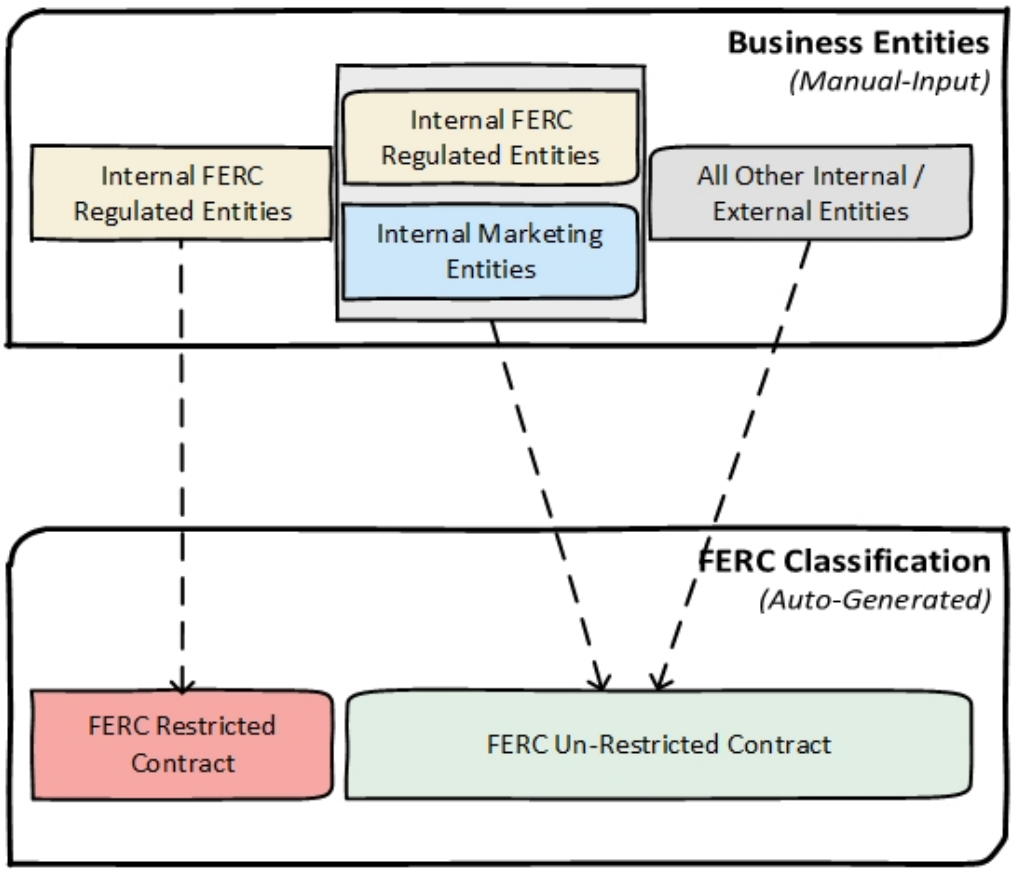

Figure 3-8 Metadata Auto-Generation Logic (FERC Regulation)

The table shown below is an example of how the FERC Classification field is auto-populated based on the metadata value manually assigned in the Internal Business Entities field.

	Sample Contract 1	Sample Contract 2
Content Function:	Legal	Legal
Content Category:	Contracts & Agreement	Contracts & Agreement
Content Type:	Commercial Contract	Commercial Contract
Content Sub-Type:	Purchase Agreement	Lease Agreement
Internal Business Entities	Company A1	Company B1
External Business Entities	Company X1	Company X1
Commodities	Natural Gas	NGL
Contract Number	CCPC-NG-00165	CCLA-NGL-00335
Execution Date	11/1/1999	5/15/2013
Termination Date	7/31/2020	12/31/2035
FERC Classification	Restricted	Un-Restricted

Table 3-15 Auxiliary Classification Metadata Mapping – Auto-Generation (Contracts)

3.2.4 MAPPING USERS TO BUSINESS FUNCTION CLASSIFICATION

In our identity management system, we need to classify employees with business function classifications. Let's assume that we have seven employees: Alice, Bob, Charlie, David, Eve, Fred, and Gary.

- Alice works under the Commercial organization in the Crude Management business unit.
- Bob works under the Commercial organization in the NGL Management business unit.
- Charlie works under the Commercial organization in the Natural Gas Management business unit.
- David and Eve works in the Legal organization.
- Fred works under the Accounting organization in the Volume Accounting business unit.
- Gary works under the Information Technology organization in the Enterprise Business Systems business unit.

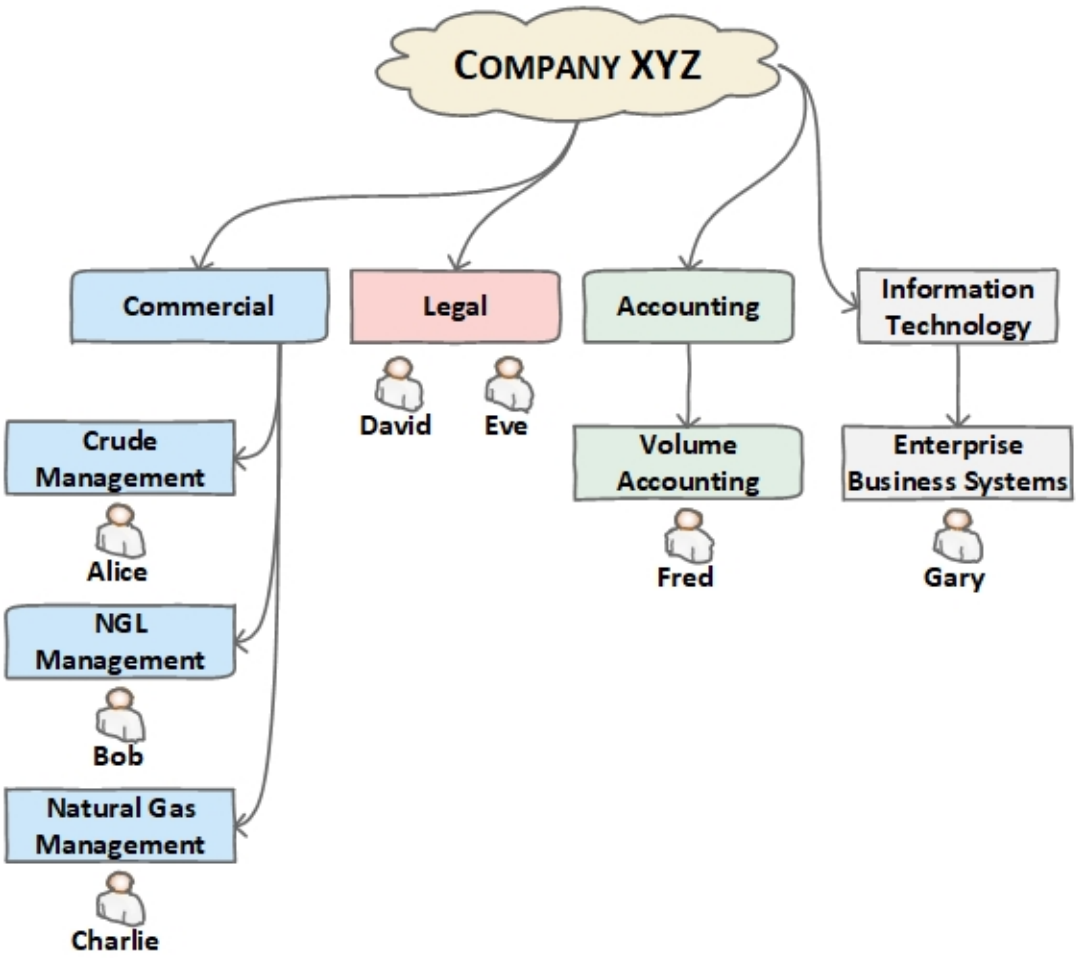

Figure 3-9 Business Functional Organization Chart

The table shown below demonstrates how the business function metadata attributes will be used to tag an employee.

Employee	Organization	Business Unit
Alice	Commercial	Crude Management
Bob	Commercial	NGL Management
Charlie	Commercial	Natural Gas Management
David	Legal	
Eve	Legal	
Fred	Accounting	Volume Accounting
Gary	Information Technology	Enterprise Business Systems

Table 3-16 Business Function Classification Metadata Mapping (User Identity)

3.2.5 MAPPING USERS WITH INFORMATION SENSITIVITY CLASSIFICATION

In our identity management system, we need to classify employees with information sensitivity classifications as security classification. This classification will help us dynamically create security rules to protect the integrity of the content.

In this security model, the classification does not directly grant read or read/write permission to any content. This classification will only serves as an attribute for create dynamic security rules. Similar to how information sensitivity classification is applied at the content level, the classification at the user level have certain predefined logic. In our example, this attribute will be used to help define the maximum level of read access permitted, regardless of the content type. Combined with other metadata attributes, dynamic rules can be created to further refine the read or read write account assignment.

Let's continue with editing the classification for seven employees. The table shown below, demonstrates information sensitivity metadata attributes are used to tag an employee.

Employee	Security Classification
Alice	Confidential
Bob	Confidential
Charlie	Sensitive
David	Confidential
Eve	Confidential
Fred	Sensitive
Gary	Internal Use Only

Table 3-17 Security Classification Metadata Mapping (User Identity)

3.2.6 MAPPING USERS TO BUSINESS ROLE CLASSIFICATIONS

In our example, employees in both Commercial and Information Technology organization handles contracts and have a need to share these files with employees within Accounting. The entire Legal organization also needs access to view all contracts and agreements. Since this is not a role-based security model, the roles assigned to an identity does not directly impact the security access control. Instead, it will only serve as another attribute for define dynamic security rules.

- Individual Role Requirements:
    - Alice, Bob, and Gary have a business role as “Contract Administrators” for their organizations.
    - Charlie, Fred, and Gary have a business role as “Contract Consumer” for their organizations.
- Organizational Role Requirements:
    - The entire Legal organization have a business role as “Contract Consumer”

The table shown below, demonstrates how business role metadata attributes are used to tag an employee.

Employee	Business Role
Alice	Contract Administrators
Bob	Contract Administrators
Charlie	Contract Consumers
David	Contract Consumers
Gary	Contract Administrators

Table 3-18 Business Role Classification Metadata Mapping (User Identity)

Table 3-20, as shown below, demonstrates how business role metadata attributes are used to tag an organization's identity.

Organization	Business Role
Legal	Contract Consumers

Table 3-19 Business Role Classification Metadata Mapping (Organization Identity)

3.2.7 MAPPING USERS TO AUXILIARY CLASSIFICATIONS

In our example, employees in the Commercial organization need to be identified with a FERC status classification. This attributes will be used in our examples to create dynamic security rules in a later section.

- Alice and Bob have a FERC status of "Regulated" to manage Regulated assets.
- Charlie has a FERC status of "Marketing" as employee with a marketing role.
- David, Eve, Fred, and Gary do not have a FERC status, as they are not part of the Commercial organization.

The table shown below demonstrates how auxiliary metadata attributes are used to tag an employees' identity.

Employee	FERC Status
Alice	Regulated
Bob	Regulated
Charlie	Marketing
David	n/a
Eve	n/a
Fred	n/a
Gary	n/a

Table 3-20 Auxiliary Classification Metadata Mapping (User Identity)

3.3 DYNAMIC SECURITY ENGINE

Now that all our prerequisites are in place, we can transition into the discussion on the Dynamic Security Engine (DSE). DSE is a middle tier service that has two major responsibilities.

- To allow system administrators to create and configure dynamic (metadata-driven) security rules
- Systematically execute the dynamic security rules based on the priority order

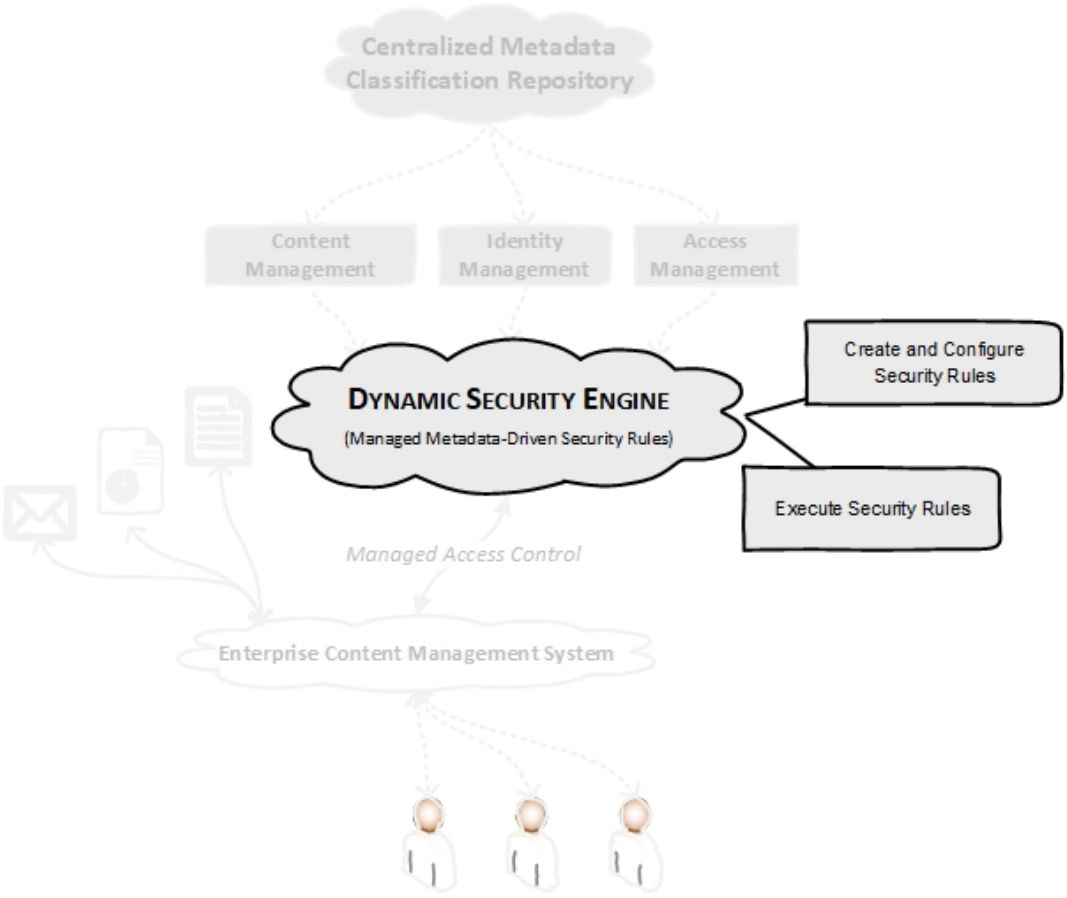

Figure 3-10 Dynamic Security Engine Overview

3.3.1 CREATING DYNAMIC SECURITY RULES

This section will explain how to leverage metadata attributes from ECM systems and attributes from IdM systems to create a dynamic security rules. Each security rule can have multiple access accounts assigned to it.

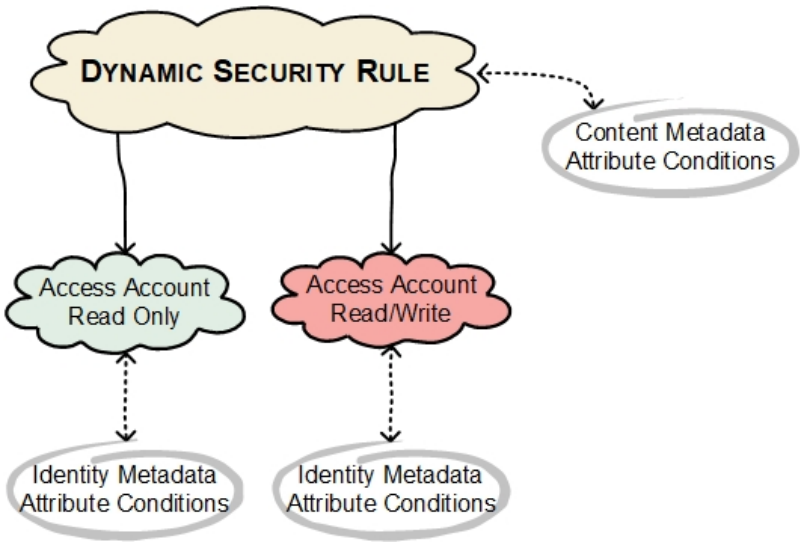

Figure 3-11 Dynamic Security Rule to Access Accounts Diagram

There are five major steps or components to creating a dynamic security rule and properly assigning access controls.

1) Create Security Rule
2) Set Security Rule's Metadata Condition with Content Attributes
3) Create the Access Accounts for Security Rule
4) Set Access Accounts Permission
5) Set Access Account's Metadata Condition with Identity Attributes

The next section will include examples of creating dynamic security rules in pseudo-codes. Each of these examples will build onto the previous and increase in complexity.

3.3.1.1 SECURITY RULE – EXAMPLE 1

REQUIREMENTS

The Legal organization needs read-only access to retrieve all contracts and agreements.

SECURITY RULE CONFIGURATION

1) Create Security Rule
    - CREATE Security_Rule = *"Contracts_All"*
2) Set Security Rule's Metadata Condition with Content Attributes
    - FOR RULE *"Contracts_All"*
        - IDENTIFY

          Content_Function = "Legal"

          Content_Category = "Contracts & Agreement"
3) Create the Access Accounts for Security Rule
    - FOR RULE *"Contracts_All"*
        - CREATE Access_Account = *"Contracts_All_Legal_R"*
4) Set Access Account's Permission
    - FOR RULE *"Contracts_All"*
        - FOR ACCOUNT *"All_Legal_R"*
            - SET PERMISSION = *READ*
5) Set Access Account's Metadata Condition with Identity Attributes
    - FOR RULE *"Contracts_All"*
        - FOR ACCOUNT *"All_Legal_R"*

- IDENTIFY

  Employee_Organization = "Legal"

SECURITY RULE OUTPUT

Dynamic Security Rule	Dynamic Access Account	Permission	Employee
Contracts_All	All_Legal_R	Read	David Eve

Table 3-21 Dynamic Security Rule Output (Example 1)

SECURITY RULE HIGHLIGHTS

Since this "Contracts_All" security rule used a metadata condition that stops at the content category level, the Legal organization will be able to see both Commercial Contracts and IT Contracts.

3.3.1.2 SECURITY RULE – EXAMPLE 2

REQUIREMENTS

The Accounting organization needs its Contract Consumers within the Volume Accounting organization to have read-only access to all Commercial Contracts.

SECURITY RULE CONFIGURATION

1) Create Security Rule
    - CREATE Security_Rule = *"CommercialContracts"*
2) Set Security Rule's Metadata Condition with Content Attributes
    - FOR RULE *"CommercialContracts"*
        - IDENTIFY

          Content_Function = "Legal"

          Content_Category = "Contracts & Agreement"

          Content_Type = "Commercial Contract"
3) Create the Access Accounts for Security Rule
    - FOR RULE *"CommercialContracts"*
        - CREATE Access_Account = *"Volume_Accounting_R"*
4) Set Access Accounts Permission
    - FOR RULE *"CommercialContracts"*
        - FOR ACCOUNT *"Volume_Accounting_R"*
            - SET PERMISSION = *READ*
5) Set Access Account's Metadata Condition with Identity Attributes
    - FOR RULE *"CommercialContracts"*

- FOR ACCOUNT *"Volume_Accounting_R"*
  - IDENTIFY

    Employee_Organization = "Accounting"

    Employee_Business_Unit = "Volume Accounting"

SECURITY RULE OUTPUT

Dynamic Security Rule	Dynamic Access Account	Permission	Employee
CommercialContracts	Volume_Accounting_R	Read	Fred

Table 3-22 Dynamic Security Rule Output (Example 2)

3.3.1.3 SECURITY RULE – EXAMPLE 3

REQUIREMENTS

The Commercial organization needs its Contract Consumers to have read-only access to all Commercial Contracts, and Contract Administrators to have read/write access to all Commercial Contract.

SECURITY RULE CONFIGURATION

Since this is rule still applies at the Contracts Content Type level (Commercial Contracts), we can skip steps 1-2.

~~1) Create Security Rule~~

~~2) Set Security Rule's Metadata Condition with Content Attributes~~

3) Create the Access Accounts for Security Rule
    - FOR RULE *"CommercialContracts"*
        - CREATE Access_Account = *"Commercial_CtrConsumer_R"*
        - CREATE Access_Account = *"Commerical_CtrAdmin_RW"*
4) Set Access Accounts Permission
    - FOR RULE *"CommercialContracts"*
        - FOR ACCOUNT *"Commercial_CtrConsumer_R"*
            - SET PERMISSION = *READ*
        - FOR ACCOUNT *"Commerical_CtrAdmin_RW"*
            - SET PERMISSION = *READ/WRITE*
5) Set Access Account's Metadata Condition with Identity Attributes
    - FOR RULE *"CommercialContracts"*

- FOR ACCOUNT *"Commercial_CtrConsumer_R"*
  - IDENTIFY

    Employee_Organization = "Commercial"

    Employee_Business_Role = "Contract Consumer"
- FOR ACCOUNT *"Commercial_CtrAdmin_RW"*
  - IDENTIFY

    Employee_Organization = "Commercial"

    Employee_Business_Role = "Contract Administrator"

SECURITY RULE OUTPUT

Dynamic Security Rule	Dynamic Access Account	Permission	Employee
CommercialContracts	Volume_Accounting_R	Read	Fred
CommercialContracts	Commercial_CtrConsumer_R	Read	Charlie
CommercialContracts	Commercial_CtrAdmin_RW	Read/Write	Alice Bob

CHAPTER 4

4 DISCUSSIONS ON IMPLEMENTATION PROCESS AND OUTCOME

4.1 HISTORICAL STUDIES

In the previous chapters, we have reviewed the specifics of the concept of the metadata-driven security model. Now, we will assess and see if there are historical attempts of using the "Attribute Based Access Control" (ABAC) security model for managing security in Enterprise Content Management, in which we specifically covered in this thesis. This in-depth research regarding ABAC will further dive into the historic events of how ABAC became recently important. By understanding the historical events, we can see a historical timeline and the evolution of ABAC. Seeing how it has grown over time will give us the insight to its potential to support this thesis topic, as a metadata-driven security model for enterprise content management systems.

As we have learned in chapter 2, Metadata-driven information security model leverages the ABAC access control model. According to the following publication by National Institute of Standards and Technology (NIST) (NIST, 2016), this except will shed light into the brief history of what initiated the need and desire of the ABAC model:

> *"The concept of Attribute Based Access Control (ABAC) has existed for many years. It represents a point on the spectrum of logical access control from simple access control lists to more capable role-based access, and finally to a highly flexible method for providing access based on the evaluation of attributes" (NIST, 2016)*
>
> *"In November 2009, the Federal Chief Information Officers Council (Federal CIO Council) published the Federal Identity, Credential, and Access Management (FICAM) Roadmap and Implementation Plan v1.0, which provided guidance to federal organizations to evolve their logical access control architectures to include the evaluation of attributes as a way to enable access within and between organizations across the Federal enterprise. In December 2011, the FICAM Roadmap and Implementation Plan v2.0 took the next step of calling out ABAC as a recommended access control model for promoting information sharing between diverse and disparate organizations" (NIST, 2016)*

From the excerpt above, we now understand that even though the concept of ABAC existed for many years, the model did not take form until November 2009. It was then, that the need to truly define this model was initiated by the Federal CIO Council. (NIST, 2016) This allowed the ABAC model to finally take form and gain maturity. The following excerpt, from the same article, provides us with a complete definition of ABAC and how the model works. This definition will sound very familiar, as it will relate to this thesis in supporting how a metadata-driven security model can be implemented with ECM.

> *"ABAC is a logical access control model that is distinguishable because it controls access to objects by evaluating rules against the*

> *attributes of the entities (subject and object) actions and the environment relevant to a request. Attributes may be considered characteristics of anything that may be defined and to which a value may be assigned. In its most basic form, ABAC relies upon the evaluation of attributes of the subject, attributes of the object, environment conditions, and a formal relationship or access control rule defining the allowable operations for subject-object attribute and environment condition combinations. All ABAC solutions contain these basic core capabilities to evaluate attributes and environment conditions, and enforce rules or relationships between those attributes and environment conditions." (NIST, 2016)*

Understating what the ABAC model comprises of along with some of its history, we will now examined how it gained popularity and what makes it unique. This is also the part of history that helps us define and demonstrate how the dynamic security engine, as discussed in chapter 3, plays its role in the technical implementation aspect of this security model. From the following excerpt below, we will learn that the greatest strength of ABAC is the flexibly it holds, in which allows an infinite amount of combinations and possibilities to defining security rules for ensuring confidentially and integrity. This flexibility is the key to why and how a metadata-driven security model is possible both conceptually and in application.

> *"The rules or policies that can be implemented in an ABAC model are limited only to the degree imposed by the computational language. This flexibility enables the greatest breadth of subjects to access the greatest breadth of objects without specifying individual relationships between each subject and each object. For example, a subject is assigned a set of subject attributes upon employment (e.g., Nancy Smith is a Nurse*

> *Practitioner in the Cardiology Department.). An object is assigned its object attributes upon creation (e.g., a folder with Medical Records of Heart Patients). Resources Objects may receive their attributes either directly from the creator or as a result of automated scanning tools. The administrator or owner of an object owner creates an access control rule to govern the set of allowable operations (e.g., all Nurse Practitioners in the Cardiology Department can View the Medical Records of Heart Patients). Adding to the flexibility of the logical access control model attributes and their values may then be modified throughout the lifecycle of subjects, objects, and attributes without modifying each and every subject/object relationship. This provides a more dynamic access control capability as access decisions can change between requests when attribute values change." (NIST, 2016)*

By creating metadata-driven security rules using attributes from a centralized managed metadata repository, these rules can be written by only applying the business knowledge and requirements of how the content should be protected or shared with whom, regardless of knowing the actual users or seeing the actual files. These security rules will provide and assign security dynamically though the dynamic security engine to ECM. From the following excerpt below, NIST provides us with both the standard definition and a business scenario to support this argument.

> *"Provisioning ABAC describes attributes to subjects and objects governed by an access control rule set that specifies what operations can take place, this capability enables object owners or administrators to apply access control policy without prior knowledge of the specific subject and for an unlimited number of subjects that might require access. As new subjects join the organization, rules and objects do not*

> *need to be modified. As long as the subject is assigned the attributes necessary for access to the required objects (e.g., all Nurse Practitioners in the Cardiology Department are assigned those attributes), no modifications to existing rules or object attributes are required. This benefit is often referred to as accommodating the external user and is one of the primary benefits of employing ABAC." (NIST, 2016)*

The concept of ABAC is refreshing and easy to understand, however it is extremely hard to define and outline exactly how it can be applicable in all the different industries. According to the same excerpt, over the past decade a few firms have tried to integrate ABAC into their security models, but all have hit roadblocks. (NIST, 2016) In my opinion, however, their failures can be corrected by defining a specific model for a specific business requirement that can leverage ABAC, such as this thesis which emphasis on applying a metadata driven security model specifically for ECM only. A copy of the referenced excerpt is below:

> *"Over the past decade, vendors have begun implementing Attribute Based Access Control (ABAC)-like features in their security management and network operating system products, without general agreement as to what constitutes an appropriate set of ABAC features. Due to a lack of consensus on ABAC features, users cannot accurately assess the benefits and challenges associated with ABAC." (NIST, 2016)*

Since the publication of the article referenced above by NIST, Cybersecurity Center of Excellence (NCCoE) hosted a workshop to discuss NIST SP 800-162, a conceptual guide for

ABAC. (NIST, 2016) “About 100 people attended the event; individuals were invited from a variety of different areas: government, industries, and academia. The workshop provided attendees an opportunity to identify, refine, and guide the many interrelated considerations, challenges, and efforts needed to develop ABAC guidance. In the workshop, three topics: 1), “The importance of ABAC”, 2), “NIST SP 800-162: Guide to Attribute Based Access Control Definition and Considerations”, and 3) “Framework of ABAC models” were presented. In the panel section, experts focused on the issues of motivation, applications, and vision of ABAC. In the demonstrations and poster display section, 13 vendors demonstrated ABAC related products and research.” (NIST, 2016) The workshop marks the start of the attempt to push for ABAC to be rapidly recognized by industry and security professionals.

Deriving from the workshop, it wasn’t until a few months ago, on October 1, 2017 that the first book, consolidating the knowledge from hundreds of research papers regarding ABAC, was officially published by the National Institute of Standards and Technology (NIST) and authored by Vincent C. Hu, David F. Ferraiolo, Ramaswamy Chandramouli, and D. Richard Kuhn. (NIST, 2016) By formally publishing this guide, NIST will surely attract more vendors to not only be aware of ABAC’s abilities, but also leverage its concept and build other models, such as this thesis, to adopt it as the future of access control.

4.2 SOFTWARE COMPANIES ADOPTING THE ABAC MODEL

A few of the vendors mentioned in the previous section continues to use ABAC, as of April, 2018. Below are three examples of software companies that offer the ABAC model as a product or solution.

- CiperPoint Software
- Elasticsearch
- Axiomatics Federal, Inc.

CiperPoint Software

On October 29, 2015, a news journal published by Document Strategy Media stated that "most recently, CipherPoint added Attribute Based Access Control (ABAC) that allows organizations to dynamically apply different access decisions according to user (subject) and information (object) metadata. The control also includes the ability to filter library contents such that users only see the existence of files for which they are able to open. This functionality is in addition to the capabilities CipherPoint has always had to restrict content views up to and including the IT Administrative staff." (Document Strategy Media, 2015)

Elasticsearch

On January 12, 2018, Elasticsearch published an article on their blog announcing that "Thanks to a new feature in Lucene 7.1, the CoveringQuery, and the exposure of that feature in the new terms_set query released with Elasticsearch 6.1, it is possible to setup an attribute-based access control (ABAC) scheme for documents stored in Elasticsearch." (Barretta, 2018)

Axiomatics Federal, Inc.

On March 12, 2018, Axiomatics published a video blog titled "Moving Beyond Identity-Based Access Control." In the article, it states "Axiomatics' solutions implement Attribute Based Access Control (ABAC). This means that, when processing an access control request, digital rules can be evaluated to consider the attributes of the requestor, the attributes of the requested resource, and the attributes of the environment. This combination of attributes lends context to

the access control request and is where the powerful advantages of ABAC really shine. For example, in a Role Based Access Control (RBAC) model, a user identity is required for making static access control decisions. While ABAC often incorporates user attributes in making access control decisions, identities are not necessarily required. With ABAC, run-time access control decisions can be made based on existing environment variables." (Colaluca, 2018)

From the above references, we can conclude that the infusion of ABAC model has been applicable and achievable for enterprises. The success of the few vendors in the market in an encouraging notion that proves the feasibly of other similar models, such as the metadata-driven security model for enterprise content management, as demonstrated in this thesis.

4.3 VALUE PROPOSITION

In any organization, regardless of the type or size, maintaining proper access control for each and every employee is a never-ending process. Based on an organization with approximately 10,000 employees, the daily activities to monitor, grant, and revoke access request to ensure business continuity can cost up to 4-5 full-time employees to support. Additionally, the risk of providing excessive access control to an employee due to departmental reorganizations, role changes, or promotions is a problem that lingers in our society. The bigger the company, the more it becomes problematic to keep up.

In the recent years, companies and software has shifted towards using role-based access control models (RBAC) in hopes of improving their ability manage and support complex business and system authorization requirements. However, RBAC has its limitations.

The following list sourced from Elasticsearch explains the key limitations.

- “The number of roles would balloon as an organization grew and as the number of kinds of data grew, making their management unwieldy” (Barretta, 2018)
- “Keeping roles mutually exclusive and collectively exhaustive (MECE) is hard, meaning it'd be possible to grant someone a number of contradictory roles that could result in a leak of data or other unintended behaviors” (Barretta, 2018)
- “Roles are meant to be generic and applicable to many people: they don't take user-specific information into account” (Barretta, 2018)

Using the attribute-based access control (ABAC) model can eliminate these limitations. Below is an excerpt extracted from a journal published by Elasticsearch, in which explains the value of ABAC.

> *Attribute-based access control depends on attributes assigned to users, things, and actions, and a policy to make decisions based on them. For a user, attributes could include projects they work on, team memberships, certifications, years of service, and physical location. For a thing (i.e. resource), attributes could be sensitivity level, PII status, time-to-live (TTL), or physical location.*
>
> *A real-world control policy that is easier to model in ABAC is printing information in a secured environment: you can only print (action) from a specific printer (resource) if you are allowed to print things (action attribute + user attribute), that printer is near your workspace (resource attribute + user attribute), and your security training is up to date (contextual information: current date + user attribute). In RBAC, you'd need a printing role, a role for each printer (how many thousands in a big org?), need to update printing role membership everyday as users slipped out of training compliance, and update the membership in*

> *the printers' roles each day as people joined, left, and moved throughout the organization. (Barretta, 2018)*

4.4 BUSINESS PROCESS AND MANAGEMENT DEMANDS

As with all business programs, management support is crucial its success. In order to use a metadata-driven information security model, the prerequisite is to already have a mature Information Management program in place. The implementation process can span across each and every organization within the company. It will require dedication and commitments some every group for the purpose of ensuring and enforcing integrity and confidentiality for all business records. Therefore, without Executive Management's support, a project this size is destined to fail. Steering committee members for sustaining this program will also need to be carefully selected. Every member in the steering committee should have a commitment to collectively establish and support a standardized metadata classification framework that is applicable for the entire company. As a society, most of us are known to be resistance to change. This security model not only requires changes from the management's vision, but also requires all users to understand how to manage content with metadata. Providing education on metadata, user training, and establishing awareness campaign is another key to success.

CHAPTER 5

5 CONCLUSION

Throughout this thesis, we started with the scope to determine if it is possible to leverage the use of metadata in Enterprise Content Management (ECM) systems to dynamically manage access control, without compromising confidentiality, and maintaining the integrity of official business records. We performed literature reviews and covered the concepts of information security, information management, enterprise content management, identity and access management, and metadata taxonomy. We also reviewed the entire concept of a metadata-driven security model designed specifically for ECM by creating sample metadata taxonomy, sample metadata values, mapping rules, and creating sample dynamic security rules with pseudo-logics. We also had discussion on implementation and outcomes with historical studies and industry example. In conclusion, we can answer the question that thesis was scoped for, confirming that it is absolutely possible to leverage the use of metadata in Enterprise Content Management (ECM) systems to dynamically manage access control, without compromising confidentiality, and maintaining the integrity of official business records.

With the recent rise in security awareness and multiple scandals of security breaches of big name corporations, the focus had been about protecting assets and information from external threats. However, a key component in information security that is critical but is often given little credit, is managing content security within an enterprise organization. It is information security

at its core, in which defines how content, data, and information is secured internally, managed, and collaborated upon among shareholders and employees. It is more important now than ever, to revisit how we manage information security, as digital content is growing exponentially worldwide.

The exponential increase in digital data has introduced the demand for enterprise content management systems, which focus on centralizing unstructured information assets into a single repository, and utilizes metadata to support dynamic search and retrieval. As such, if the enterprise data architecture is designed properly, this metadata-driven security model (ABAC model) can be utilized to dynamically manage access control and protect sensitive material through the use of metadata classifications. Although this model is only at the stage of infancy, this model seems to be gaining traction as demand for information sharing and risk of information security breaches are both on the rise. This model will require a strong information management program, along with various technical components to be implemented together. It will also require a sustainable change management workflow to support the ever-changing information taxonomy, as the business evolves over time. Information taxonomy and metadata management is critical and must be properly defined, managed, and aligned with compliance standards to provide a truly sustainable metadata-driven security model for ECM.

In conclusion, this thesis was to research and answer the question "*Is it possible to leverage the use of metadata in Enterprise Content Management (ECM) systems to dynamically manage access control, without compromising confidentiality, and maintaining the integrity of official business records?"* The short and simple answer is, "Yes." In fact, should an organization choose to implement this this security model, they will surely benefit from the high level of integrity and confidential provided by it. The outcome of run-time access control

management, with automatic user account allocation based on real-time business decisions, will definitely reduce risk and cost for a company or organization.

REFERENCES

AIIM. (2018). *What is Electronic Records Management (ERM)?* Retrieved from AIIM.

AIIM. (2018, April 18). *What is Enterprise Content Management (ECM)?* Retrieved from AIIM: http://www.aiim.org/What-is-ECM-Enterprise-Content-Management#

AIIM. (2018, March 6). *What is Information Management*. Retrieved March 6, 2018

Akkermans, D. (2018, April 18). *Institute of Morphoid Research*. Retrieved from Taxonomy Chart: http://imr.jenniferakkermans.com/2011/06/taxonomy-chart/

Alfresco Software Inc. (2018). *Alfresco Software Inc.* Retrieved from The File Plan: https://docs.alfresco.com/rm/concepts/rm-fileplan.html

Allen, T. (2011, 01 10). *The Importance of Metadata in Content Management* . Retrieved 12 04, 2017, from CMSWiRE: https://www.cmswire.com/cms/enterprise-cms/the-importance-of-metadata-in-content-management-009746.php

Anitha, R. a. (2016, 05 12). METADATA DRIVEN EFFICIENT KEY GENERATION AND DISTRIBUTION IN CLOUD SECURITY. *Journal of Computer Science 10 (10): 1929-1938, 2014*, 1929.

Araujo, C. (2017, 10 16). *How a Metadata-Driven Architecture Creates Organizational Agility* . Retrieved 12 04, 2017, from INTELLYX: https://intellyx.com/2017/10/16/how-a-metadata-driven-architecture-creates-organizational-agility/

Ball, J. (2013, 09 30). *NSA stores metadata of millions of web users for up to a year, secret files show*. Retrieved 12 03, 2017, from https://www.theguardian.com/world/2013/sep/30/nsa-americans-metadata-year-documents

Barretta, M. (2018, January 12). *Document-Level Attribute-Based Access Control with X-Pack 6.1*. Retrieved from Elasticsearch: Document-Level Attribute-Based Access Control with X-Pack 6.1

Barroca, E. (2016, 10 20). *Content management platforms can unlock the full value of digital assets* . Retrieved 12 04, 2017, from GCN: https://gcn.com/articles/2016/10/20/content-management-platforms.aspx

Basic Biology. (2017, April 18). *Taxonomy*. Retrieved from Basic Biology: https://basicbiology.net/biology-101/taxonomy/

Berners-Lee, T. (1989, March). *W3C.* Retrieved April 18, 2018, from Information Management: A Proposal: https://www.w3.org/History/1989/proposal.html

Cain, A. (2017, 10 31). *Taxonomy*. (Encyclopædia Britannica, inc.) Retrieved 03 07, 2018, from Encyclopædia Britannica: https://www.britannica.com/science/taxonomy

CERN. (2018, April 19). *Tim Berners-Lee's Proposal*. Retrieved from CERN: http://info.cern.ch/Proposal.html

Charles Crichton, J. D. (2007, 12 01). *Metadata-Driven Software for Clinical Trials*. Retrieved 12 04, 2017, from Oxford University Computing Laboratory: http://www.cs.ox.ac.uk/people/jeremy.gibbons/publications/consort.pdf

CIMTECH. (2007). *Managing Information and Documents*. Retrieved from Hacettepe University: http://www.bby.hacettepe.edu.tr/akademik/ozgurkulcu/file/Managing%20Documents%20and%20Record,ERM_%C3%96nemli%20Kitap.pdf

Colaluca, B. (2018, March 12). *Axiomatics*. Retrieved from Video Blog-Moving Beyond Identity-Based Access Control: https://www.axiomatics.com/blog/video-blog-moving-beyond-identity-based-access-control/

Comentum. (2010, 11 15). *What is a Content Management System (CMS)?* Retrieved from Comentum: http://www.comentum.com/what-is-cms-content-management-system.html

Conklin, W. A. (2014). Introduction to Incident Response. Texas.

Conklin, W. A., White, G., Williams, D., Davis, R., & Cothren, C. (2016). *Principles of Computer Security, Fourth Edition.* McGraw-Hill.

David Schlesinger, C. (2010, 08 01). *Taxonomy to Drive Data Quality, Security and Regulatory Compliance in a Global Enterprise*. Retrieved 12 04, 2017, from The Software Engineering Institute (SEI), Carnegie Mellon University: https://www.sei.cmu.edu/measurement/research/upload/Schlesinger.pdf

Document Strategy Media. (2015, October 29). *DocPoint & CipherPoint Enhancing Data Security within Government Agencies' SharePoint ECM Deployments*. Retrieved from Document Strategy: http://documentmedia.com/article-2298-DocPoint-CipherPoint-Enhancing-Data-Security-within-Government-Agencies'-SharePoint-ECM-Deployments.html

Garbarini, A. (2012, 01 01). *Data Lake for the Enterprise*. Retrieved 12 04, 2017, from Big Data Everywhere: http://www.bigdataeverywhere.com/files/chicago-2016/BDE-DataLakePresentation-GARBARINI.pdf

Gartner. (2018). *Identity and Access Management (IAM)*. Retrieved from Gartner: https://www.gartner.com/it-glossary/identity-and-access-management-iam/

Gartner, Inc. (n.d.). *Information Governance*. Retrieved March 14, 2018, from Gartner.: https://www.gartner.com/it-glossary/information-governance/

Harvard Law School. (2018). *What is Metadata?* Retrieved April 10, 2018, from Harvard Law School: https://hls.harvard.edu/dept/its/what-is-metadata/

Khan Academy. (2018, April 18). *Cuneiform*. Retrieved from Khan Academy: https://www.khanacademy.org/humanities/ancient-art-civilizations/ancient-near-east1/sumerian/a/cuneiform

Laminin Solutions. (2013, 02 19). *Metadata Permissions Protects Confidential Information* . Retrieved 12 04, 2017, from Laminin Solutions: https://lamininsolutions.com/2013/02/19/metadata-permissions-protects-confidential-information/

Masson, S. (2016, 06 22). *FIVE REASONS WHY CLASSIFICATION IS THE FIRST STEP TO SUCCESSFUL DATA-LOSS PREVENTION*. Retrieved 12 04, 2017, from Data Center Journal. All Right Reserved. Designed and Developed by Renew Design Studios: http://www.datacenterjournal.com/five-reasons-classification-first-step-successful-data-loss-prevention/

Mukherjee, R. A. (2014). METADATA DRIVEN EFFICIENT KEY GENERATION AND DISTRIBUTION IN CLOUD SECURITY . *Journal of Computer Science* , 1929-1938.

National Institute of Standards and Technology. (1992, October 13). *Role-Based Access Controls.* (D. R. David F. Ferraiolo, Ed.) Retrieved April 14, 2018, from National Institute of Standards and Technology: https://csrc.nist.gov/CSRC/media/Publications/conference-paper/1992/10/13/role-based-access-controls/documents/ferraiolo-kuhn-92.pdf

NIST. (2016, May 24). *Attribute Based Access Control*. Retrieved from National Institude of Standards and Technology: https://csrc.nist.gov/projects/abac/

Obama, B. (2009, December 29). *Executive Order 13526—Classified National Security Information.* Retrieved from U.S. Department of State Freedom of Information Act: https://foia.state.gov/_docs/MDR/135190.pdf

Office of the Director of National Intelligence. (2017, 01 01). *Information Security Marking Metadata*. Retrieved 12 04, 2017, from Office of the Director of National Intelligence: https://www.dni.gov/index.php/who-we-are/organizations/ic-cio/ic-cio-related-menus/ic-cio-related-links/ic-technical-specifications/information-security-marking-metadata

PaperFree Corporation. (2018). *Enterprise Content Management (ECM)*. Retrieved from PaperFree Corporation: http://www.paperfreecorp.com/ecm

Philips, H. (2010, September). *The Great Library of Alexandria?* Retrieved April 10, 2018, from University of Nebraska - Lincoln: https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1427&context=libphilprac

Privacy International. (1990, 01 01). *Metadata*. Retrieved 12 04, 2017, from Privacy International: https://www.privacyinternational.org/node/53

Raggad, B. G. (2010). *Information Security Management: Concepts and Practice* . Baca Raton, FL: CRC Press.

Riley, J. (2017, January 1). *Understanding Metadata: What is Metadata, and What is it For?: A Primer.* Retrieved April 10, 2018, from NISO: http://groups.niso.org/apps/group_public/download.php/17446/Understanding%20Metadata.pdf

Seiner, R. S. (2003, 04 01). *A Conceptual Meta-Model for Unstructured Data* . Retrieved 12 04, 2017, from The Data Administration Newsletter, LLC: http://tdan.com/a-conceptual-meta-model-for-unstructured-data/5087

Subramanian, P. (2015, 01 01). *Security Content Metadata Model with an Efficient Search Methodology for Real Time Monitoring and Threat Intelligence*. Retrieved 12 04, 2017, from Semantic Scholar: https://pdfs.semanticscholar.org/0611/edd2070372f097748257f395e92c24eef452.pdf

The Official Website of the Executive Office for Administration and Finance. (2014, 03 06). *Enterprise Information Security Standards: Data Classification* . Retrieved 12 04, 2017, from Commonwealth of Massachusetts: http://www.mass.gov/anf/research-and-tech/cyber-security/security-for-state-employees/security-policies-and-standards/enterprise-information-security-standards.html

University of Illinois. (2018, April 16). *Data Classifications*. Retrieved from University of Illinois at Chicago Academic Computing and Communications Center: https://security.uic.edu/data-classifications/

White, R. H. (2009, 12 01). *Writing security policies using a taxonomy-based approach*. Retrieved 12 04, 2017, from SearchSecurity: http://searchsecurity.techtarget.com/magazineContent/Writing-security-policies-using-a-taxonomy-based-approach

Williams, B. D. (2017, 07 28). *Researcher: Metadata the 'most potent weapon' against critical infrastructure security*. Retrieved 12 04, 2017, from Federal Times: https://www.federaltimes.com/critical-infrastructure/2017/07/28/researcher-metadata-the-most-potent-weapon-against-critical-infrastructure-security/

Young, J. (2016, 03 21). *Mediachain: Protect Digital Content With a Bitcoin-Based Metadata Protocol*. Retrieved 12 04, 2017, from BITCOINMAGAZINE: https://bitcoinmagazine.com/articles/mediachain-protect-digital-content-with-a-bitcoin-based-metadata-protocol-1458577763/