An XML Schema for Electronic Records Management

DRAFT
Version 0.1

by

Leila Naslavsky
and
Dorrit Gordon

1 Introduction

1.1 Background/Overview

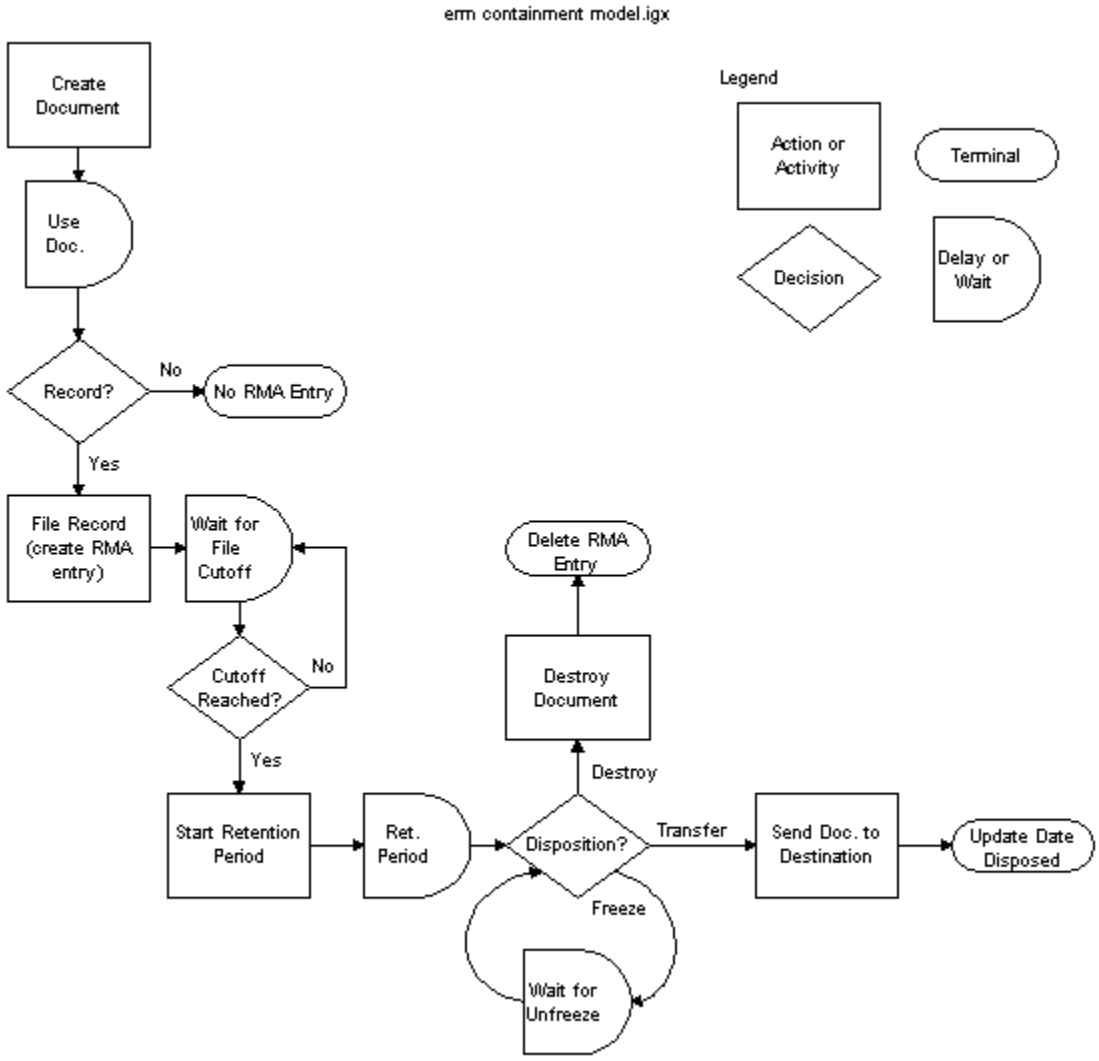
Governments, corporations, museums, libraries and other organizations all maintain repositories of documents. As the volume of information maintained by these institutions grows, so does the need for simple, flexible mechanisms for managing that information. Electronic records management (ERM) is a field of study devoted to developing methods for managing this information.

This work defines an XML schema to be used as the basis for a records management application (RMA) compliant with the United States Department of Defense's (DoD) standards for RMAs, DoD 5015.2 [3].

The United States government defines records as any "books, papers, maps, photographs, machine readable materials, or other documentary materials, regardless of physical form or characteristics, made or received by an agency of the United States Government under Federal law or in connection with the transaction of public business and preserved or appropriate for preservation by that agency or its legitimate successor as evidence of the organization, functions, policies, decisions, procedures, operations, or other activities of the Government or because of the informational value of data in them. Library and museum material made or acquired and preserved solely for reference or exhibition purposes, extra copies of documents preserved only for convenience of reference, and stocks of publications and of processed documents are not included." [11]

All documents handled by the DoD follow a similar lifecycle. At some point after a document is created, a decision is made regarding its status as a record. If it is determined to be a record, the document is filed. To file a record is to create an entry for it in the RMA, and to assign it to at least one file. Records are also assigned to a record category which, among other things, associates with the record a disposition instruction. Files are periodically cutoff (meaning no more records may be added to them). When a file is cutoff, all records belonging to that file enter their retention periods. Retention periods are defined in a record's associated disposition instruction. At the end of its retention period, a record is disposed of according to the disposition action defined in it's associated disposition instruction. Temporary records are destroyed and permanent records are transferred to another office, or to an archive site. Occasionally a record must be retained beyond the end of its normal retention period. A disposition action code of 'freeze' indicates this. Frozen records are retained until their disposition action code is changed to some other value. Figure 1 illustrates the record lifecycle.

Figure 1 Record Lifecycle



1.2 Schema Design Goals

The XML schema defined in this paper was designed with two key goals in mind: compliance with DoD 5015.2, and flexibility.

We chose to use DoD 5015.2 as a basis for our schema. It provides a manageable set of clearly defined metadata elements. The US based ERM research has been driven by the National Archive and Records Administration (NARA) which relies on the DoD 5015.2 standard. An Australian organization, the Public Record Office of Victoria has also published a comprehensive set of ERM standards, the Victorian Electronic Records Strategy (VERS). The VERS Metadata Scheme [9] is so large as to be outside the scope of this project.

In addition to compliance with DoD 5015.2, we designed our schema to be as flexible as possible. We chose to implement the schema in several pieces to allow organizations which

structure records differently from the DoD and other branches of the US government to utilize only the elements of the schema which are appropriate for their organizations. Our inclusion of facilities for user defined fields both supports a DoD 5015.2 requirement, and allows other organizations to define additional metadata as desired. There are two limiting factors in our schema.

The first is the identification of required and optional fields. Our schema follows the DoD 5015.2 requirements. This will force other organizations using our schema to comply with DoD 5015.2 metadata requirements. An implementation based on our schema can easily overcome this restriction. For fields which the DoD standard defines as 'required', implementers that wish them to be absent (or optional) can select a default value to assign them if no is other specified by the user. An implementation can also require fields that the DoD standard defines as optional by screening the user's input and rejecting any that is missing values for those fields.

The second, and more significant, limitation is our adherence to the DoD specification's rules on the number of occurrences of a given field. Fields which this schema limits to a single occurrence per entity, cannot be altered by implementation to allow multiple occurrences per entity. An organization requiring multiple occurrences per entity would need to use user-defined data in place of the standard element included in this schema.

2 Data Model

This section introduces the first- and second-class data elements which make up our ERM system. The first-class data elements are those which form the structure of the system. The second-class elements are those which hold individual items of information.

2.1 Major Data Elements

Based on the DoD standard, we selected seven first-class data elements. These are defined in Table 1.

Table 1 First-Class Data Elements

Name	Definition
Record	An RMA entry which contains metadata pertaining to a particular document (or other work) which must by managed in the RMA.
File	An RMA entry which contains records, and which has an associated File Plan which defines how to manage those records.
File Plan	An RMA entry which defines how records in a particular file must be managed.
Record Category	An RMA entry used to associate sets of records with a particular disposition instruction.
Disposition Instruction	An RMA entry which defines the lifespan of records with which it is associated, and which identifies what should be done with records when their active lifespan is over.
File Category	An RMA entry used to associate sets of files with a particular cutoff instruction.

Name	Definition
Cutoff Instruction	An RMA entry which defines the frequency with which a file should be cutoff. When a file is cutoff, the current instance is closed and a new instance is opened. Records in the instance being closed begin their retention periods as per their disposition instructions.

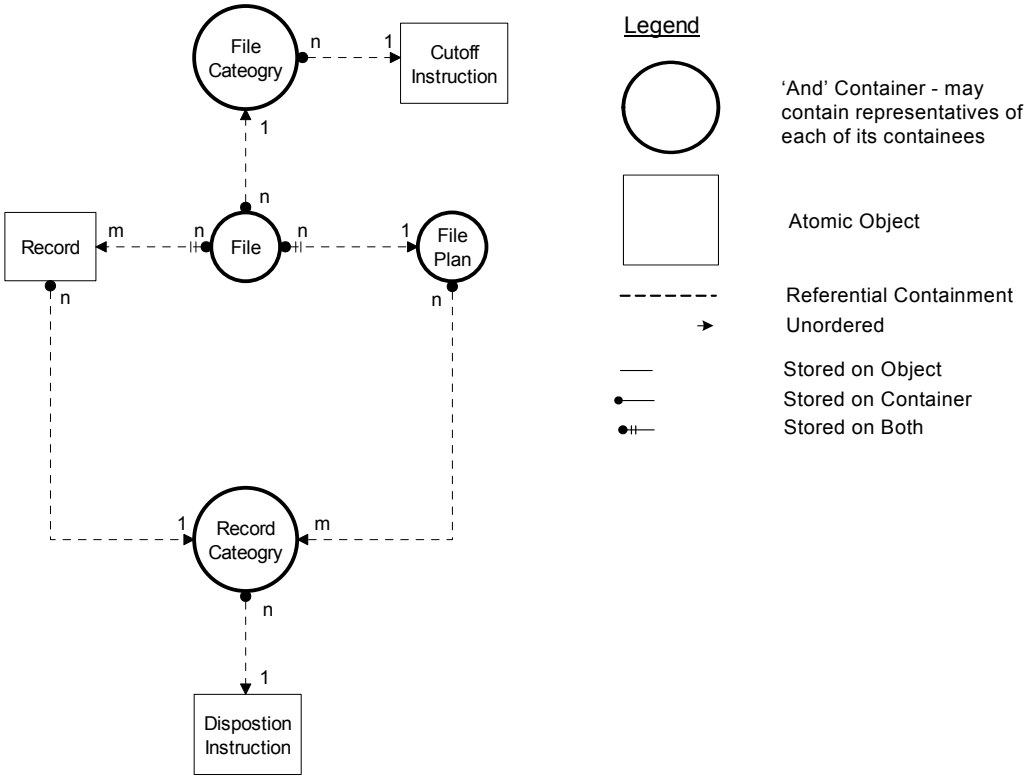
Since records are the heart of any ERM system, they were clearly a first-class object. According to the DoD specification, files also have important properties in their own right, so we selected them as another first-class object. Our analysis of DoD 5015.2 leads us to believe that file plans apply to multiple files. Had there been a one-to-one relationship between file plans and files, then file plans could be considered second-class data elements, subsidiary to files. The one-to-many relationship, however, requires that they be made independent, first-class objects. Similar arguments apply to record categories with respect to records, and to file categories with respect to files.

The biggest question surrounds disposition instructions and cutoff instructions. The primary purpose of record categories is to associate records with disposition instructions. Since each record category identifies only one disposition instruction, it would have been reasonable to make disposition instructions secondary to record categories. We elected to make disposition instructions first-class options for several reasons. First, DoD 5015.2 seems to allow a disposition instruction to apply to multiple record categories. To avoid duplication of data, it made sense to make disposition instructions independent data elements. Second, a disposition instruction is a conceptually independent, stand-alone unit of information. Subordinating it to some other data element would mask the importance of the concept. Similar relationships exist between files, file categories and cutoff instructions, and similar arguments apply.

2.2 Containment Model

Figure 2 shows the containment structure for a DoD 5015.2 compliant records management application. A record is the basic element of any RMA system. In the Department of Defense, each record belongs to one or more files and to one or more record categories within those files. Each record category has an associated disposition instruction which defines what should happen to records in that category at the end of their active life spans. Each file contains a file plan which defines the record categories available for that file, as well as other general information related to the file. Furthermore, each file is categorized. The file category associates a cutoff instruction with the file. The cutoff instruction defines how frequently a file instance is closed, and a new instance opened.

Figure 2 DoD 5015.2 Containment Model



2.3 Detailed Metadata

This section describes the second-class data elements. The section is organized by first-class element. The names of the data elements correspond to the names of the elements in the XML schema.

2.3.1 Record

Records are the heart of an ERM system. Most of the elements we chose to include for records are explicitly identified in DoD 5015.2. The only exception is the 'predecessor' field. DoD 5015.2 explicitly refers to a 'successor' field which holds an identifier for a record which supercedes this one. It seemed to us that a user might also want to find a preceding version of a record. For example, if an agency is investigating a potential violation of one of its standards, the investigators would want to review the version of the standard in effect at the time of the alleged violation, not necessarily the current version. The ability to access a predecessor (or chain of predecessors) of a record will make this search easier. Clearly, in heavily versioned systems where access of old versions is common, a more sophisticated mechanism is necessary, but this mechanism should suffice for most records management applications.

ID	R1

Element Name	subject
Description	Allows users to enter a short subject line (similar to an email message's subject line) for a record.
Type	string
Mandatory	yes
Repeatable	no
Comments	

ID	R2
Element Name	fileDate
Description	The date the record was filed.
Type	date
Mandatory	yes
Repeatable	no
Comments	This field should be automatically set by the interface application.

ID	R3
Element Name	transmissionReceiptData
Description	complex type (see below)
Type	transmissionReceiptType
Mandatory	no
Repeatable	no
Comments	transmissionReceiptType is defined in this schema. It may contain either one emailData element or one correspondenceData element. emailData must contain at least one addressee (of type string,) and may contain one or more otherRecipients (of type string). Both addressee and otherRecipient elements may have attributes containing an email address. correspondenceData elements must contain at least one addressee (of type string) and may contain one or more otherRecipients (of type string).

ID	R4
Element Name	location
Description	This field is used to describe the physical (or electronic) location of the record.
Type	string
Mandatory	yes
Repeatable	no
Comments	For a future draft, the authors should consider constructing a complex data type for this field. The complex type would include several options: freeform entry (as is the case in this draft), a URL (with pattern matching to force proper URL construction), and perhaps selection from a set of option defined the organizational schema (for organizations with a fixed set of possible record locations, or with a fixed syntax for expressing locations).

ID	R5

Element Name	format
Description	This field is used to describe the format of the record. For electronic records this will be the file type and version (e.g. Word97, Visio 5.0). For paper records, it may be as simple as 'paper' or it may be more specific (e.g. Form X) if there is some special format that applies to the document.
Type	string
Mandatory	yes
Repeatable	no
Comments	

ID	R6
Element Name	media
Description	This field is used to identify the type of media on which the record is stored.
Type	mediaType
Mandatory	yes
Repeatable	no
Comments	Available values for must be defined by the organization using this schema.

ID	R7
Element Name	creationDate
Description	Indicates the date the document was created (may be different from the date it was filed).
Type	date
Mandatory	yes
Repeatable	no
Comments	

ID	R8
Element Name	originator
Description	Contains identifying information for the document's originator (or author).
Type	originatorType
Mandatory	yes
Repeatable	yes
Comments	originatorType is defined in this schema. This type offers a choice of four elements: personalName, officialTitle, officeSymbol, code. personalName is a string. The other three elements are to be defined by the organizations which use this schema. A personalName is intended to identify an individual (e.g. Joe Smith). An officialTitle is intended to identify a titled office (e.g. Secretary of Defense) rather than the person currently filling the office. officeSymbol and code are other ways of identifying the titled office.

ID	R9
Element Name	originatingOrg
Description	Contains identifying information for the organization which originated the document.

Type	originatingOrgType
Mandatory	yes
Repeatable	yes
Comments	originatingOrgType is defined in this schema. This type offers a choice of two elements: officialName or code. Possible values for each are defined by the organizations which use this schema. The officialName identifies an organization by its formal name (e.g. Department of Defense). code offers an alternative method of identification.

ID	R10
Element Name	vitalRecord
Description	This field indicates whether the record is considered to be vital.
Type	boolean
Mandatory	yes
Repeatable	no
Comments	The default value for this field is 0 (i.e. *not* vital).

ID	R11
Element Name	fileCode
Description	This field associates this record with a particular file or files.
Type	unsignedInt
Mandatory	yes
Repeatable	yes
Comments	An application which implements this schema must ensure that values entered identify files which actually exist in the system. This element may have a NULL value.

ID	R12
Element Name	recordCateogryCode
Description	This field associates this record with a particular record category or categories.
Type	unsignedInt
Mandatory	yes
Repeatable	no
Comments	An application which implements this schema must ensure that values entered identify record categories which actually exist in the system and which are valid categories for at least one file with which this record is associated. This element may have a NULL value.

ID	R13
Element Name	ID
Description	This field is a unique identifier for the record.
Type	recordIDType
Mandatory	yes
Repeatable	no
Comments	An application which implements this schema must autogenerate this value.

	The value must be a 10 digit number prefixed by a 4-digit calendar year. These numbers must be unique across the system. They should be assigned sequentially.

ID	R14
Element Name	userDefinedData
Description	This field allows users to add their own data.
Type	userDefinedDataType
Mandatory	no
Repeatable	yes
Comments	userDefinedDataType is defined in this schema. It contains two values: dataFieldName and dataFieldValue. Both elements are strings. They are used to hold name for the field and a value for the field respectively.

ID	R15
Element Name	relatedRecords
Description	This field allows users to indicate other records which are associated with this record.
Type	recordIDType
Mandatory	no
Repeatable	yes
Comments	See type definition for ID. For the purposes of this schema, we made an assumption that it is unnecessary to store metadata about related information which is not part of the ERM system.

ID	R16
Element Name	successor
Description	This field is used to identify a record which has replaced this record as the most current version of the document.
Type	recordIDType
Mandatory	no
Repeatable	no
Comments	See type definition for ID. If this field has a value, it must refer to an existing record.

ID	R17
Element Name	predecessor
Description	This field is used to identify a record which is replaced by this record.
Type	recordIDType
Mandatory	no
Repeatable	no
Comments	See type definition for ID. If this field has a value, it must refer to an existing record.

ID	R18

Element Name	dispositionActionDate
Description	This field is used to mark the date on which a record should be disposed.
Type	date
Mandatory	yes
Repeatable	no
Comments	The value for this element can be calculated by adding the retention period identified in the Disposition Instruction which pertains to this record to the cutoff date which applies to the file to which this record belongs (note: for files with periodic cutoff, use the next cutoff date after the record is filed).

2.3.2 Disposition Instruction

Disposition instructions describe what is to happen to a record once its active lifespan is complete. The actual content of a disposition instruction is included in the ERM system. Disposition instructions define a retention period, and the action to be taken at the end of that period.

With the exception of the code and retentionPeriod fields, all fields in the disposition instruction type are explicitly referenced by DoD 5015.2 as part of a disposition instruction. The code field we added simply as a way of uniquely identifying disposition instructions. Retention periods are mentioned several times in the DoD specification, although it never clearly stated where they should belong. We chose to include them as part of the disposition instruction because conceptually retention periods seem to be related to the disposition of documents. Effectively a disposition can be thought of as keeping the record for a certain amount of time, then destroying (or transferring as appropriate).

ID	DI1
Element Name	name
Description	This field gives a human-readable identifier for the disposition instruction.
Type	string
Mandatory	yes
Repeatable	no
Comments	

ID	DI2
Element Name	code
Description	This field contains an identifier for the disposition instruction.
Type	unsignedInt
Mandatory	yes
Repeatable	no
Comments	This field must be autogenerated by any application which implements the schema. Values must be unique across the system. Numbers should be assigned sequentially.

ID	DI3

Element Name	type
Description	This field identifies a type of disposition. Possibilities include time, event, or event-time.
Type	complex type (see comment below)
Mandatory	yes
Repeatable	no
Comments	This field has three, fixed value choices: time, event, and time-event.

ID	DI4
Element Name	actionType
Description	This field identifies the type of action which should be taken when a record reaches its disposition time. Options are 'transfer', 'destroy', or 'freeze'.
Type	complex type (see comment below)
Mandatory	yes
Repeatable	no
Comments	This element has three fixed value choices: transfer, destroy, or freeze.

ID	DI5
Element Name	retentionPeriod
Description	This field identifies how long a record should be kept after the file is cutoff.
Type	short
Mandatory	yes
Repeatable	no
Comments	This element has an attribute, unit, which must be one of the following: day, week, month, year.

ID	DI6
Element Name	eventDescription
Description	This field can be used to include a description of the event portion of an event or event-time based disposition.
Type	string
Mandatory	no
Repeatable	no
Comments	Although not mandatory, this field should be used when type=event or type=event-time.

ID	DI7
Element Name	actionDescription
Description	This field can be used to include a description of the disposition action to be used for this record.
Type	string
Mandatory	no
Repeatable	no
Comments	Although not mandatory, this field should be used to provide information about where to transfer the record when actionCode=transfer. It may also be used to

	provide further information regarding other actions (e.g. "Freeze until completion of XYZ lawsuit").

2.3.3 File

Files group records into conceptually related collections. Records may belong to multiple files. The actual content of files is stored in the system, in the sense that the metadata for a file contains identifiers for all the records in that file, and an identifier for the file plan associated with the file.

We added cutoffStartDate and closureDate to the basic set of metadata described for files in DoD5015.2. The whole issue of cutoffs is not very clearly described in the specification. In order to support the concept in a reasonable way it seems necessary to indicate when cutoffs should begin, and when a file should be permanently closed. No metadata related to cutoffs was explicitly described in DoD 5015.2. Definition AP1.1.9 in the DoD specification reads, "[t]o cutoff records in a file means to break, or end them at regular intervals…" According to this definition, cutoffs occur on a periodic basis. Although it is not explicitly stated, we concluded that the cutoff periods may not start on the date the file is created. The cutoffStartDate field allows the users to indicate when to start measuring cutoff periods. Furthermore, the definition does not elucidate file closure. The difference between periodic cutoff and permanent closure is not defined, but we inferred that some files may close permanently (without starting a new instance of that file). For example, a presidential correspondence file might be cutoff every two months to allow old correspondence to be archived. If the file is tied to the particular person serving as president, then when he or she leaves office, the file will be permanently closed.

ID	F1
Element Name	filePlanCode
Description	This field is used to identify the file plan which defines this file.
Type	unsignedInt
Mandatory	yes
Repeatable	no
Comments	Any application which implements this schema must ensure that this value identifies a file plan which exists in the system.

ID	F2
Element Name	code
Description	This field is used to identify this file.
Type	unsignedInt
Mandatory	yes
Repeatable	no
Comments	This field must be automatically assigned by any application which implements this schema. Values must be unique across the system. Numbers should be assigned sequentially.

ID	F3
Element Name	recordsFiled

Description	This field is used to identify records which this file contains.
Type	recordID
Mandatory	no
Repeatable	yes
Comments	See type description for the ID element in Record.

ID	F4
Element Name	cutoffStartDate
Description	This field is used to set the date on which this file's cutoff cycle should begin.
Type	date
Mandatory	no
Repeatable	no
Comments	The value for this field should default to the date the file is created, but the user should have the option to change it.

ID	F5
Element Name	closureDate
Description	This field is used to indicate the date on which this file should be permanently closed (i.e. no more records should be added, no more cutoffs should be performed, and disposition instructions should be applied to all remaining records).
Type	date
Mandatory	no
Repeatable	no
Comments	This field is no mandatory because it may not always be possible to foresee the total lifetime of a file when that file is created. Some files will be permanently active, or at least will be active until some, unspecified and unforeseen event occurs which dictates closing the file. For example, when a new agency is created, it will start a file for personnel records. If the agency is a permanent one, there will be no foreseeable time at which the personnel file should be closed. If, however, the government decides at some future time to shut down the agency, then its personnel file should be closed.

ID	F6
Element Name	fileCategoryCode
Description	This field identifies the file category to which this file belongs.
Type	unsignedInt
Mandatory	yes
Repeatable	no
Comments	

2.3.4 File Plan

File plans define the record categories which are valid for a file, and identify the disposition authority document which formally grants permission to the appropriate people to dispose of records in that file.

ID	FP1
Element Name	code
Description	This field is to identify this file plan.
Type	unsignedInt
Mandatory	yes
Repeatable	no
Comments	This field must be automatically assigned by any application which implements this schema. Values must be unique across the system. Numbers should be assigned sequentially.

ID	FP2
Element Name	recordCateogryCode
Description	This field is used to identify record categories which are valid for the file defined by this file plan.
Type	unsignedInt
Mandatory	yes
Repeatable	yes
Comments	Any application which implements this schema must ensure that values for this field identify record categories which are defined in the system.

ID	FP3
Element Name	dispositionAuthority
Description	This field is used to identify the disposition authority document associated with the file defined by this file plan.
Type	string
Mandatory	yes
Repeatable	no
Comments	

ID	FP4
Element Name	userDefinedData
Description	This field allows system users to add additional data as desired.
Type	userDefinedDataType
Mandatory	no
Repeatable	yes
Comments	See type description for userDefinedData in Record.

2.3.5 Record Category

Record categories associate records with disposition instructions. All elements included here were explicitly described in 5015.2.

ID	RC1
Element Name	name
Description	This field provides a human-readable identifier for this record category.
Type	string
Mandatory	yes
Repeatable	no
Comments	

ID	RC2
Element Name	code
Description	This field is used to identify this record category.
Type	unsignedInt
Mandatory	yes
Repeatable	no
Comments	This field must be automatically assigned by any application which implements this schema. Values must be unique across the system. Numbers should be assigned sequentially.

ID	RC3
Element Name	description
Description	This field allows users to provide a description of this record category.
Type	string
Mandatory	yes
Repeatable	no
Comments	

ID	RC4
Element Name	dispositionInstructionCode
Description	This field is used identify the disposition instruction associated with this record category.
Type	unsignedInt
Mandatory	yes
Repeatable	no
Comments	Any application which implements this schema must ensure that the value of this field identifies a disposition instruction which exists in the system.

ID	RC5
Element Name	userDefinedData
Description	This field allows users to add their own data.
Type	userDefinedDataType
Mandatory	no

Repeatable	yes
Comments	See userDefinedData in Record for a description of this type.

2.3.6 File Category

File categories associate files with cutoff instructions.

File categories are only obliquely referred to in DoD 5015.2. When we decided that cutoff instructions should be a first-class object, we also decided the files and records should have parallel support structures. Therefore we added file categories to serve the same role for files that record categories serve for records. The content of file category is parallel to that of record category.

ID	FC1
Element Name	name
Description	This field provides a human-readable identifier for this file category.
Type	string
Mandatory	yes
Repeatable	no
Comments	

ID	FC2
Element Name	code
Description	This field is used to identify this file category.
Type	unsignedInt
Mandatory	yes
Repeatable	no
Comments	This field must be automatically assigned by any application which implements this schema. Values must be unique across the system. Numbers should be assigned sequentially.

ID	FC3
Element Name	description
Description	This field allows users to provide a description of this file category.
Type	string
Mandatory	yes
Repeatable	no
Comments	

ID	FC4
Element Name	cutoffInstructionCode
Description	This field is used identify the cutoff instruction associated with this file category.
Type	unsignedInt
Mandatory	yes

Repeatable	no
Comments	Any application which implements this schema must ensure that the value of this field identifies a cutoff instruction which exists in the system.

ID	FC5
Element Name	userDefinedData
Description	This field allows users to add their own data.
Type	userDefinedDataType
Mandatory	no
Repeatable	yes
Comments	See userDefinedData in Record for a description of this type.

2.3.7 Cutoff Instruction

Cutoff instructions specify the frequency with which files are cutoff.

As mentioned in 2.3.3, cutoffs are not clearly described in DoD 5015.2. This element represents an effort to provide a mechanism which meets the limited definition available. The definition provided in section AP1.1.9 in the DoD specification describes periodic cutoff. We were also able to think of cases where a file might be cutoff only once, at the end of its lifetime. For example a file created to hold records related to a court case probably would not be cutoff periodically, instead it would be cutoff a single time, at file closure. We decided to use a simple type flag to identify which case applied, and to include a field to hold the frequency for periodic cutoffs. We also questioned what should happen in cases where a cutoff cannot be performed on the date scheduled (for example, because the date falls on a non-business day). Although the specification does not answer this question, we elected to provide a comment field which could be used to provide instructions (or for any other purpose).

ID	IC1
Element Name	code
Description	This field identifies the cutoff instruction.
Type	unsignedInt
Mandatory	yes
Repeatable	no
Comments	

ID	IC2
Element Name	comment
Description	This field allows users to enter textual information about the cutoff instruction.
Type	string
Mandatory	no
Repeatable	no
Comments	One suggested use for this field is to contain information describing appropriate actions for cutoff dates which fall on non-business days. For example, should the cutoff activities be performed on the preceding business day? The next business day? Must they be performed exactly on the right date regardless of

	day?

ID	IC3
Element Name	cutoffType
Description	This field defines the frequency with which cutoffs should occur for files to which this cutoff instruction applies.
Type	complex type (see comment below)
Mandatory	yes
Repeatable	no
Comments	This element may either have the fixed value, 'single,' meaning that the file has only one cutoff, or it may give a frequency for cutoffs. Frequencies are of type, erm:periodType which is of type short. erm:periodType also has an attribute which must take a value of: day, week, month or year.

2.3.8 Example Organization Definitions

This section provides an example of the schema an organization might use to define values that our schema leaves open. The assumption is that any organization which uses our full set of schemas, will take this example and modify it as appropriate for their organization. They will define the official titles, office names, office symbols and office codes which are used in their organization. They may elect to use as complete a list of mediaTypes as they can come up with. The mediaType field could also be used to limit the kinds of media acceptable for the storage of records.

The decision to make some aspects of the schema definable by individual organizations has the potential to limit interoperability between implementations. Many organizations have very specific sets of values which are appropriate for the data elements included in this schema. We decided that supporting the need to limit input to acceptable values was more important in the short term than interoperability. In early use of these systems, it is unlikely that organizations will share records. So long as future implementations are backwards compatible, there should be no barrier to developing more interoperable system in the future. Furthermore, if there is a need for sharing records between organizations, this limitation will provide incentive for organizations to come together to develop joint standards for these metadata elements. A consensus standard will be a strong basis for future development of this schema.

ID	OS1
Element Name	mediaType
Description	This element defines the possible types of media on which a record may be stored.
Type	string
Mandatory	yes
Repeatable	no
Comments	Possible values: CD-R DAT DVD

	floppy disk hard disk JAZ drive microfilm paper videotape WORM ZIP drive audio cassette CD-rewriteable other This list derives from [9].

ID	OS2
Element Name	officialTitle
Description	This elements identifies the official job titles available in the organization.
Type	string
Mandatory	yes
Repeatable	no
Comments	Possible values: Peon Pointy Haired Boss Big Boss Top Brass

ID	OS3
Element Name	officeSymbol
Description	This element contains the possible office symbols available in this organization.
Type	string
Mandatory	yes
Repeatable	no
Comments	Possible values: DoD DoE DoT DoI

ID	OS4
Element Name	code
Description	This element defines the office codes used by an organization.
Type	string
Mandatory	yes
Repeatable	no
Comments	Possible values: Dept1 Dept2

	DeptEtc

ID	OS5
Element Name	officialName
Description	This element defines the official office names available in an organization.
Type	string
Mandatory	yes
Repeatable	no
Comments	Possible values: Department of Defense Department of Energy Department of Transportation Department of Education Department of the Interior

3 Issues Addressed by Implementation

There are some requirements in DoD 5015.2 which cannot be enforced by XML. These requirements must be enforced by an application which implements the schema.

Perhaps the most important of these requirements is an implicit requirement that records must be associated with existing files and record categories, and the files be associated with existing file plans and file categories. Currently this schema does not provide a mechanism for ensuring that the codes used to associate records with files and record categories identify files and record categories for which entries have been created. An implementation of the schema would have to enforce this requirement by performing some consistency check whenever a record is created. There are similar consistency issues surrounding the codes used to associate record categories with file plans, and disposition instructions with record categories.

A second requirement which can be enforced in implementation, but not in the schema is that records for which the vital record indicator is set to 'yes' are not disposed of. The schema cannot enforce any limitations on actions relating to an element (e.g. deleting or updating some other field) based on the value of a field. An application which implements this schema must enforce this requirements.

Finally, the schema cannot enforce consistency or accuracy in any other references (e.g. to the document's author, or to the disposition authority). Any consistency or accuracy checking must be performed by the application or by the RMA's administrators.

4 Conclusion

4.1 General Conclusion

The volume of recorded information is growing exponentially. As the volume of information grows, tools for managing that information must grow in size and in sophistication. The schema defined in this document provides a basis for such tools.

The primary goal of this effort was to support the US Department of Defense's requirements for electronic records management. To that end, the metadata included in our schema derives directory from DoD 5015.2. We added a limited number of additional metadata elements. Elements were added either because their presence was inferred from the specification (although not explicitly defined), or because we felt they would add significant utility. Elements which were added only because we felt they would add utility are all optional.

The secondary goal of this effort was to create a schema with enough flexibility to be used for other ERM efforts. We chose to split the schema into several pieces. An organization should use only those pieces which are appropriate to its needs. Records are the core of any RMA. Our assumption is that any organization implementing our schema will, at a minimum, use the records portion. The only additional portion required is the organizational schema. The File Schema cannot be implemented without the Record Schema because files contain elements with a type, recordID, which is defined in the Record Schema. All other interconnections are via elements of type unsignedInt. These can be implemented with meaningless default values, when they do not refer to the code for an existing element. The inclusion of user defined data elements allows organizations to include metadata that is not called for in the DoD 5015.2 specification.

While there is still considerable work to be done to ensure that this schema is fully compliant with DoD 5015.2, and to improve its flexibility and support for interoperability, this draft will provide a platform from which to move forward.

4.2 Related Work

There has been considerable research into the field of electronic record preservation [2]. The research deals with methods for ensuring that data stored electronically remains available for use when needed. Issues include degradation of electronic media, and obsolescence of electronic data formats. Our work is less concerned with the physical storage of records, than with maintaining the metadata necessary to identify, locate, and track records (electronic or otherwise). To ensure the long-term availability of records both lines of research are crucial. The ability to locate records is useless if, once located, the records are unreadable due to degradation of the physical storage medium, or to storage formats so out of date that tools are no longer available to read them. Likewise, it is of no use to have readable data if it is impossible to identify which data is of use.

As mentioned several times in this paper, VERS is an alternate standard for electronic records management. Ideally VERS, DoD 5015.2 and any other ERM metadata standards will evolve to a compatible state. The reconciliation of multiple standards was outside the scope of this effort. Further, because these standards derive from existing, non-electronic management systems, there is no guarantee that the competing processes are fully compatible.

The International Electrotechnical Commission (IEC) has drafted a metadata standard for technical documents [7]. The notion of defining metadata for documents is similar to the ERM metadata effort, but the ISO effort does have significant differences from our effort. First, the work refers to a particular type of document (technical document). Second, the document lifecycle described in [7] is quite different from the record lifecycle inferred from DoD 5015.2. Thus, the metadata requirements in the two standards are rather different. That being said, the

IEC standard should certainly be considered in any effort to produce a truly general-purpose ERM schema.

4.3 Future Work

This work has not been verified to be in compliance with DoD 5015.2. It will need to be reviewed by outside reviewers to ensure compliance with the standard. In particular, this schema may not yet be in full compliance with respect to email messages filed as records.

In the interests of flexibility, and interoperability, other standards must be investigated in more detail. This schema may need revision to better support a variety of standards. One current barrier to interoperability is the Organization Schema. We need to define a mechanism that will allow organizations to limit the values available for the elements described in that schema, without inhibiting interoperability.

This schema focuses on a government standard for records management. Many non-government agencies also maintain and manage records. There has been no investigation into the needs of these other agencies.

Records management is related to archival science. The relationship of this work to archival sciences must be explored.

5 Schema

5.1 Record Schema

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- Record Profile Schema-->
<schema targetNamespace="http://electronicrecordsmanagement" xmlns:orgnamespace="http://organizationXMLSchema"
xmlns:erm="http://electronicrecordsmanagement" xmlns="http://www.w3.org/2001/XMLSchema" elementFormDefault="unqualified"
attributeFormDefault="unqualified">
    <import namespace="http://organizationXMLSchema" schemaLocation="organizationXMLSchema.xsd"/>
    <simpleType name="recordIDType">
        <annotation>
            <documentation>Functional Baseline Requirements and Data Elements for Records Management Application Software
- 3.10 RECORD_Identifier </documentation>
            <documentation>An application which implements this schema must autogenerate this value. The value must be a 10
digit number prefixed by a 4-digit calendar year. These numbers must be unique across the system. They should be assigned
sequentially.</documentation>
        </annotation>
        <restriction base="unsignedLong">
            <pattern value="[1-9]{1}[0-9]{13}"/>
        </restriction>
    </simpleType>
    <complexType name="userDefinedDataType">
        <annotation>
            <documentation>DoD 5015.2 STD-C2.2.2.7.</documentation>
        </annotation>
        <sequence>
            <element name="dataFieldName" type="string"/>
            <element name="dataFieldValue" type="string"/>
        </sequence>
    </complexType>
    <complexType name="originatingOrgType">
        <annotation>
            <documentation>DoD 5015.2 STD-AP1.1.41.</documentation>
        </annotation>
        <choice>
            <element name="officialName" type="orgnamespace:officialName"/>
```

```
            <element name="code" type="orgnamespace:code"/>
        </choice>
    </complexType>
    <complexType name="originatorType">
        <annotation>
            <documentation>DoD 5015.2 STD-AP1.1.7.</documentation>
        </annotation>
        <choice>
            <element name="personalName" type="string"/>
            <element name="officialTitle" type="orgnamespace:officialTitle"/>
            <element name="officeSymbol" type="orgnamespace:officeSymbol"/>
            <element name="code" type="orgnamespace:code"/>
        </choice>
    </complexType>
    <simpleType name="emailAddressType">
        <annotation>
            <documentation>e-mail pattern</documentation>
        </annotation>
        <restriction base="string">
            <pattern value="([&#x21;-&#x7E; -[@,;:"\[\]\(\)\\&#x3c;&#x3E;]])@(&#x21;-&#x7E; -[@,;:"\[\]\(\)\\&#x3c;&#x3E;]])"/>
        </restriction>
    </simpleType>
    <complexType name="TransmissionReceiptType">
        <annotation>
            <documentation>DoD 5015.2 STD-C2.2.3.2.</documentation>
            <documentation>Detailed Metadata R3</documentation>
        </annotation>
        <choice>
            <element name="emailData">
                <complexType>
                    <sequence>
                        <element name="addressee" maxOccurs="unbounded">
                            <complexType>
                                <simpleContent>
                                    <extension base="string">
                                        <attribute name="emailAddress" type="erm:emailAddressType" use="required"/>
                                    </extension>
                                </simpleContent>
                            </complexType>
                        </element>
                        <element name="otherRecipient" maxOccurs="unbounded">
                            <complexType>
                                <simpleContent>
                                    <extension base="string">
                                        <attribute name="emailAdress" type="erm:emailAddressType" use="required"/>
                                    </extension>
                                </simpleContent>
                            </complexType>
                        </element>
                    </sequence>
                </complexType>
            </element>
            <element name="correspondenceData">
                <annotation>
                    <documentation>National Archives and Records Administration Regulations: 36 CFR Part 1234.22 (b)
</documentation>
                </annotation>
                <complexType>
                    <sequence>
                        <element name="addressee" type="string" maxOccurs="unbounded"/>
                        <element name="otherRecipient" type="string" maxOccurs="unbounded"/>
                    </sequence>
                </complexType>
            </element>
        </choice>
    </complexType>
    <complexType name="recordType">
        <annotation>
            <documentation>DoD 5015.2 STD-AP1.1.50.</documentation>
        </annotation>
```

```
        <sequence>
            <element name="ID" type="erm:recordIDType">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.2.2.</documentation>
                    <documentation>Detailed Metadata R13</documentation>
                </annotation>
            </element>
            <element name="fileDate" type="date">
                <annotation>
                    <documentation>DoD 5015.2 TD-C2.2.2.5.2., AP1.1.15.</documentation>
                    <documentation>Detailed Metadata R2</documentation>
                </annotation>
            </element>
            <element name="TransmissionReceiptData" type="erm:TransmissionReceiptType" minOccurs="0">
                <annotation>
                    <documentation>Detailed Metadata R3</documentation>
                </annotation>
            </element>
            <element name="media" type="orgnamespace:mediaType">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.2.5.4.</documentation>
                    <documentation>Detailed Metadata R6</documentation>
                </annotation>
            </element>
            <element name="format" type="string">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.2.5.5.</documentation>
                    <documentation>Detailed Metadata R5</documentation>
                </annotation>
            </element>
            <element name="location" type="string">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.2.5.6.</documentation>
                    <documentation>Detailed Metadata R4</documentation>
                </annotation>
            </element>
            <element name="creationDate" type="date">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.2.5.7., AP1.1.14.</documentation>
                    <documentation>Detailed Metadata R7</documentation>
                </annotation>
            </element>
            <element name="originator" type="erm:originatorType" maxOccurs="unbounded">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.2.5.8., AP1.1.7.</documentation>
                    <documentation>Detailed Metadata R8</documentation>
                </annotation>
            </element>
            <element name="originatingOrg" type="erm:originatingOrgType" maxOccurs="unbounded">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.2.5.9., AP1.1.41.</documentation>
                    <documentation>Detailed Metadata R9</documentation>
                </annotation>
            </element>
            <element name="vitalRecord" type="boolean" default="0">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.1.3.5., C2.2.2.5.10., C2.2.2.12., AP1.1.59.</documentation>
                    <documentation>Detailed Metadata R10</documentation>
                </annotation>
            </element>
            <element name="fileCode" type="unsignedInt" nillable="true">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.2.1., AP1.1.32.; This element is nilable to allow the implementation of
systems that do not use the concept of record category such as VERS .</documentation>
                    <documentation>Detailed Metadata R11</documentation>
                </annotation>
            </element>
            <element name="recordCategoryCode" type="unsignedInt" nillable="true">
                <annotation>
```

```
                    <documentation>DoD 5015.2 STD-C2.2.1.3.2., AP1.1.46.; This element is nilable to allow the implementation of systems that do not use the concept of record category such as VERS .
                    </documentation>
                    <documentation>Detailed Metadata R12</documentation>
                </annotation>
            </element>
            <element name="userDefinedData" type="erm:userDefinedDataType" minOccurs="0" maxOccurs="unbounded">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.2.7.</documentation>
                    <documentation>Detailed Metadata R14</documentation>
                </annotation>
            </element>
            <element name="relatedRecords" type="erm:recordIDType" minOccurs="0" maxOccurs="unbounded">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.2.15.</documentation>
                    <documentation>Detailed Metadata R15</documentation>
                </annotation>
            </element>
            <element name="sucessor" type="erm:recordIDType" minOccurs="0">
                <annotation>
                    <documentation>DoD 5015.2 STD-C3.2.3.</documentation>
                    <documentation>Detailed Metadata R16</documentation>
                </annotation>
            </element>
            <element name="predecessor" type="erm:recordIDType" minOccurs="0">
                <annotation>
                    <documentation>DoD 5015.2 STD-C3.2.3.</documentation>
                    <documentation>Detailed Metadata R17</documentation>
                </annotation>
            </element>
            <element name="dispositionActionDate" type="date">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.5.1, AP1.1.21..</documentation>
                    <documentation>Detailed Metadata R18</documentation>
                </annotation>
            </element>
        </sequence>
        <attribute name="subject" type="string" use="required">
            <annotation>
                <documentation>DoD 5015.2 STD-C2.2.2.5.1.</documentation>
                <documentation>Detailed Metadata R1</documentation>
            </annotation>
        </attribute>
    </complexType>
</schema>
```

5.2 Disposition Instruction Schema

```
<?xml version="1.0" encoding="UTF-8"?>
<schema targetNamespace="http://electronicrecordsmanagement" xmlns:erm="http://electronicrecordsmanagement" xmlns="http://www.w3.org/2001/XMLSchema" elementFormDefault="unqualified" attributeFormDefault="unqualified">
    <complexType name="dispositionInstructionType">
        <annotation>
            <documentation>DoD 5015.2 STD-AP1.1.23. </documentation>
        </annotation>
        <sequence>
            <element name="name" type="string">
                <annotation>
                    <documentation>Detailed Metadata DI1</documentation>
                </annotation>
            </element>
            <element name="code" type="unsignedInt">
                <annotation>
                    <documentation>DoD 5015.2 STD-AP1.1.24.</documentation>
                    <documentation>Detailed Metadata DI2</documentation>
                </annotation>
            </element>
            <element name="type">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.5.2., AP1.1.25.</documentation>
```

```
                <documentation>Detailed Metadata DI3</documentation>
            </annotation>
            <simpleType>
                <restriction base="string">
                    <enumeration value="Time"/>
                    <enumeration value="Event-Time"/>
                    <enumeration value="Event"/>
                </restriction>
            </simpleType>
        </element>
        <element name="eventDescription" type="string" minOccurs="0">
            <annotation>
                <documentation>Detailed Metadata DI6</documentation>
            </annotation>
        </element>
        <element name="actionType">
            <annotation>
                <documentation>DoD 5015.2 STD-AP1.1.20.</documentation>
                <documentation>Detailed Metadata DI4</documentation>
            </annotation>
            <simpleType>
                <restriction base="string">
                    <enumeration value="Transfer"/>
                    <enumeration value="Destroy"/>
                    <enumeration value="Freeze"/>
                </restriction>
            </simpleType>
        </element>
        <element name="actionDescription" type="string" minOccurs="0">
            <annotation>
                <documentation>Detailed Metadata DI7</documentation>
            </annotation>
        </element>
        <element name="retentionPeriod" type="erm:periodType">
            <annotation>
                <documentation>DoD 5015.2 STD-AP1.1.52.</documentation>
                <documentation>Detailed Metadata DI5</documentation>
            </annotation>
        </element>
    </sequence>
</complexType>
<complexType name="periodType">
    <annotation>
        <documentation>This field represents a certain period of time</documentation>
    </annotation>
    <simpleContent>
        <extension base="short">
            <attribute name="unit">
                <simpleType>
                    <restriction base="string">
                        <enumeration value="day"/>
                        <enumeration value="week"/>
                        <enumeration value="month"/>
                        <enumeration value="year"/>
                    </restriction>
                </simpleType>
            </attribute>
        </extension>
    </simpleContent>
</complexType>
</schema>
```

5.3 File Schema

```
<?xml version="1.0" encoding="UTF-8"?>
<schema targetNamespace="http://electronicrecordsmanagement" xmlns:erm="http://electronicrecordsmanagement"
xmlns="http://www.w3.org/2001/XMLSchema" elementFormDefault="unqualified" attributeFormDefault="unqualified">
    <include schemaLocation="RecordProfileSchema.xsd"/>
    <include schemaLocation="dispositionInstructionSchema.xsd"/>
    <complexType name="filePlanType">
```

```
        <annotation>
            <documentation>DoD 5015.2 STD-C2.2.1.3.</documentation>
        </annotation>
        <sequence>
            <element name="code" type="unsignedInt">
                <annotation>
                    <documentation>This field uniquely identifies the file plan</documentation>
                    <documentation>Detailed Metadata FP1</documentation>
                </annotation>
            </element>
            <element name="recordCategoryCode" type="unsignedInt" maxOccurs="unbounded">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.1.3.2.</documentation>
                    <documentation>Detailed Metadata FP2</documentation>
                </annotation>
            </element>
            <element name="dispositionAuthority" type="string">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.1.3.4.</documentation>
                    <documentation>Detailed Metadata FP3</documentation>
                </annotation>
            </element>
            <element name="userDefinedData" type="erm:userDefinedDataType" minOccurs="0" maxOccurs="unbounded">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.1.3.9.</documentation>
                    <documentation>Detailed Metadata FP4</documentation>
                </annotation>
            </element>
        </sequence>
    </complexType>
    <complexType name="fileType">
        <sequence>
            <element name="code" type="unsignedInt">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.2.1., AP1.1.32.</documentation>
                    <documentation>Detailed Metadata F2</documentation>
                </annotation>
            </element>
            <element name="filePlanCode" type="unsignedInt">
                <annotation>
                    <documentation>This field is used to identify the file plan which defines this file.</documentation>
                    <documentation>Detailed Metadata F1</documentation>
                </annotation>
            </element>
            <element name="recordsFiled" type="erm:recordIDType" minOccurs="0" maxOccurs="unbounded">
                <annotation>
                    <documentation>DoD 5015.2 STD-AP1.1.31.1.</documentation>
                    <documentation>Detailed Metadata F3</documentation>
                </annotation>
            </element>
            <element name="fileCategoryCode" type="unsignedInt">
                <annotation>
                    <documentation>This field is a reference to the file category of the file.</documentation>
                    <documentation>Detailed Metadata F6</documentation>
                </annotation>
            </element>
            <element name="cutoffStartDate" type="date" minOccurs="0">
                <annotation>
                    <documentation>This field is the first cutoff date and the date from which the cutoff frequency is first
measured.</documentation>
                    <documentation>Detailed Metadata F4</documentation>
                </annotation>
            </element>
            <element name="closureDate" type="date" minOccurs="0">
                <annotation>
                    <documentation>This field is the date which the file is closed (no more addition of records
allowed).</documentation>
                    <documentation>Detailed Metadata F5</documentation>
                </annotation>
            </element>
```

```
        </sequence>
    </complexType>
    <complexType name="recordCategoryType">
        <annotation>
            <documentation>DoD 5015.2 STD-AP1.1.45.</documentation>
        </annotation>
        <sequence>
            <element name="code" type="unsignedInt">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.1.3.2.</documentation>
                    <documentation>Detailed Metadata RC2</documentation>
                </annotation>
            </element>
            <element name="name" type="string">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.1.3.1.</documentation>
                    <documentation>Detailed Metadata RC1</documentation>
                </annotation>
            </element>
            <element name="description" type="string">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.1.3.3.</documentation>
                    <documentation>Detailed Metadata RC3</documentation>
                </annotation>
            </element>
            <element name="dispositionInstructionCode" type="unsignedInt">
                <annotation>
                    <documentation>DoD 5015.2 STD-C2.2.1.4.</documentation>
                    <documentation>Detailed Metadata RC4</documentation>
                </annotation>
            </element>
            <element name="userDefinedData" type="erm:userDefinedDataType" minOccurs="0" maxOccurs="unbounded">
                <annotation>
                    <documentation>Detailed Metadata RC5</documentation>
                </annotation>
            </element>
        </sequence>
    </complexType>
    <complexType name="fileCategoryType">
        <annotation>
            <documentation>Analogous to recordCategoryType</documentation>
        </annotation>
        <sequence>
            <element name="code" type="unsignedInt">
                <annotation>
                    <documentation>This field uniquely identifies the file category type.</documentation>
                    <documentation>Detailed Metadata FC2</documentation>
                </annotation>
            </element>
            <element name="name" type="string">
                <annotation>
                    <documentation>Detailed Metadata FC1</documentation>
                </annotation>
            </element>
            <element name="description" type="string">
                <annotation>
                    <documentation>Detailed Metadata FC3</documentation>
                </annotation>
            </element>
            <element name="cutoffInstructionCode" type="unsignedInt">
                <annotation>
                    <documentation>This field is a reference to the cutoff instruction associated with this file
category.</documentation>
                    <documentation>Detailed Metadata FC4</documentation>
                </annotation>
            </element>
            <element name="userDefinedData" type="erm:userDefinedDataType" minOccurs="0" maxOccurs="unbounded">
                <annotation>
                    <documentation>Detailed Metadata FC5</documentation>
                </annotation>
```

```
            </element>
        </sequence>
    </complexType>
    <complexType name="cutoffInstructionType">
        <annotation>
            <documentation>DoD 5015.2 STD-AP1.1.9.</documentation>
        </annotation>
        <sequence>
            <element name="code" type="unsignedInt">
                <annotation>
                    <documentation>This field uniquely identifies the cutoff Instruction</documentation>
                    <documentation>Detailed Metadata IC1</documentation>
                </annotation>
            </element>
            <element name="comment" type="string" minOccurs="0">
                <annotation>
                    <documentation>Detailed Metadata IC2</documentation>
                </annotation>
            </element>
            <element name="cutoffType">
                <annotation>
                    <documentation>Detailed Metadata IC3</documentation>
                </annotation>
                <complexType>
                    <choice>
                        <element name="single" type="string" fixed="single"/>
                        <element name="periodic" type="erm:periodType">
                            <annotation>
                                <documentation>This field describes how often the cutoff should happen </documentation>
                            </annotation>
                        </element>
                    </choice>
                </complexType>
            </element>
        </sequence>
    </complexType>
</schema>
```

5.4 Example Organizational Schema

```
<!-- Possible example of an organization_s defined schema-->
<schema targetNamespace="http://organizationXMLSchema" xmlns="http://www.w3.org/2001/XMLSchema"
xmlns:orgnamespace="http://organizationXMLSchema" elementFormDefault="unqualified" attributeFormDefault="unqualified">
    <simpleType name="mediaType">
        <annotation>
            <documentation>Detailed Metadata OS1</documentation>
        </annotation>
        <restriction base="string">
            <enumeration value="CD-R"/>
            <enumeration value="DAT"/>
            <enumeration value="DVD"/>
            <enumeration value="floppy disk"/>
            <enumeration value="hard disk"/>
            <enumeration value="JAZ drive"/>
            <enumeration value="microfilm"/>
            <enumeration value="paper"/>
            <enumeration value="videotape"/>
            <enumeration value="WORM"/>
            <enumeration value="ZIP drive"/>
            <enumeration value="audio cassette"/>
            <enumeration value="CD-rewriteable"/>
            <enumeration value="other"/>
        </restriction>
    </simpleType>
    <simpleType name="officialTitle">
        <annotation>
            <documentation>Detailed Metadata OS2</documentation>
        </annotation>
        <restriction base="string">
            <enumeration value="Peon"/>
```

```
            <enumeration value="Pointy Haired Boss"/>
            <enumeration value="Big Boss"/>
            <enumeration value="Top Brass"/>
        </restriction>
    </simpleType>
    <simpleType name="officeSymbol">
        <annotation>
            <documentation>Detailed Metadata OS3</documentation>
        </annotation>
        <restriction base="string">
            <enumeration value="DoD"/>
            <enumeration value="DoE"/>
            <enumeration value="DoT"/>
            <enumeration value="DoI"/>
        </restriction>
    </simpleType>
    <simpleType name="code">
        <annotation>
            <documentation>Detailed Metadata OS4</documentation>
        </annotation>
        <restriction base="string">
            <enumeration value="Dept1"/>
            <enumeration value="Dept2"/>
            <enumeration value="DeptEtc"/>
        </restriction>
    </simpleType>
    <simpleType name="officialName">
        <annotation>
            <documentation>Detailed Metadata OS5</documentation>
        </annotation>
        <restriction base="string">
            <enumeration value="Department of Defense"/>
            <enumeration value="Department of Energy"/>
            <enumeration value="Department of Transportation"/>
            <enumeration value="Department of the Interior"/>
        </restriction>
    </simpleType>
    <annotation>
        <documentation>This is an illustrative schema, it's purpose is to allow for flexibility. The organization that uses the group of
schemas defined in this document, would have to customize this particular schema by providing it's own data. For example, there
might be an organization that only uses "CD-rewritable" or "ZIP drive" media types. Their "mediaType" type would contain only those
two enumerations.</documentation>
    </annotation>
</schema>
```

5.5 RMA Schema

```
<!--This schema brings all of the elements required for a DoD 5015.2-compliant RMA together.-->
<?xml version="1.0" encoding="UTF-8"?>
<schema targetNamespace="http://electronicrecordsmanagement" xmlns="http://www.w3.org/2001/XMLSchema"
xmlns:erm="http://electronicrecordsmanagement" elementFormDefault="qualified" attributeFormDefault="unqualified">
    <include schemaLocation="RecordProfileSchema.xsd"/>
    <include schemaLocation="DispositionInstructionSchema.xsd"/>
    <include schemaLocation="FileSchema.xsd"/>
    <element name="RMASystem">
        <annotation>
            <documentation>This is the schema for a RMA system</documentation>
        </annotation>
        <complexType>
            <sequence>
                <element name="record" type="erm:recordType" maxOccurs="unbounded"/>
                <element name="file" type="erm:fileType" maxOccurs="unbounded"/>
                <element name="filePlan" type="erm:filePlanType" maxOccurs="unbounded"/>
                <element name="recordCategory" type="erm:recordCategoryType" maxOccurs="unbounded"/>
                <element name="dispositionInstruction" type="erm:dispositionInstructionType" maxOccurs="unbounded"/>
                <element name="fileCategory" type="erm:fileCategoryType" maxOccurs="unbounded"/>
                <element name="cutoffInstruction" type="erm:cutoffInstructionType" maxOccurs="unbounded"/>
            </sequence>
        </complexType>
    </element>
```

```
</schema>
```

5.6 RMA Example

```
<!--This is an example of how the RMA Schema could be used.-->
<?xml version="1.0" encoding="UTF-8"?>
<RMASystem xmlns="http://electronicrecordsmanagement" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://electronicrecordsmanagement/RMASystemSchema.xsd">
    <record subject="String">
        <ID>10000000000000</ID>
        <fileDate>1967-08-13</fileDate>
        <TransmissionReceiptData>
            <emailData>
                <addressee emailAddress="A@A">String</addressee>
                <addressee emailAddress="A@A">String</addressee>
                <otherRecipient emailAdress="A@A">String</otherRecipient>
                <otherRecipient emailAdress="A@A">String</otherRecipient>
            </emailData>
        </TransmissionReceiptData>
        <media>CD-R</media>
        <format>MSWord2000</format>
        <location>http://someURL.com/document</location>
        <creationDate>1967-08-13</creationDate>
        <originator>
            <personalName>Author's name</personalName>
        </originator>
        <originator>
            <officialTitle>Big Boss</officialTitle>
        </originator>
        <originatingOrg>
            <officialName>Department of Defense</officialName>
        </originatingOrg>
        <originatingOrg>
            <code>Dept2</code>
        </originatingOrg>
        <vitalRecord>1</vitalRecord>
        <fileCode>12345</fileCode>
        <recordCategoryCode>12</recordCategoryCode>
        <userDefinedData>
            <dataFieldName>Project Number</dataFieldName>
            <dataFieldValue>3.1</dataFieldValue>
        </userDefinedData>
        <relatedRecords>10000000000001</relatedRecords>
        <sucessor>10000000000001</sucessor>
        <dispositionActionDate>1967-08-13</dispositionActionDate>
    </record>
    <record subject="String">
        <ID>10000000000001</ID>
        <fileDate>1967-08-13</fileDate>
        <TransmissionReceiptData>
            <correspondenceData>
                <addressee>Tom</addressee>
                <addressee>John</addressee>
                <otherRecipient>Luiz</otherRecipient>
            </correspondenceData>
        </TransmissionReceiptData>
        <media>paper</media>
        <format>A4 210 X 297mm</format>
        <location>Building A, 3rd floor, Room 36, file cabinel A-3</location>
        <creationDate>1967-08-13</creationDate>
        <originator>
            <sender>A@A</sender>
        </originator>
        <originator>
            <officeSymbol>DoT</officeSymbol>
        </originator>
```

```
        <originatingOrg>
            <officialName>Department of Defense</officialName>
        </originatingOrg>
        <originatingOrg>
            <officialName>Department of Transportation</officialName>
        </originatingOrg>
        <vitalRecord>1</vitalRecord>
        <fileCode>111111</fileCode>
        <recordCategoryCode>14</recordCategoryCode>
        <userDefinedData>
            <dataFieldName>Project Number</dataFieldName>
            <dataFieldValue>3.1</dataFieldValue>
        </userDefinedData>
        <relatedRecords>10000000000000</relatedRecords>
        <predecessor>10000000000000</predecessor>
        <dispositionActionDate>1967-08-13</dispositionActionDate>
    </record>
    <file>
        <code>12345</code>
        <filePlanCode>54321</filePlanCode>
        <recordsFiled>10000000000000</recordsFiled>
        <recordsFiled>10000000000001</recordsFiled>
        <fileCategotyCode>101</fileCategotyCode>
        <cutoffStartDate>1967-08-13</cutoffStartDate>
        <closureDate>1967-08-13</closureDate>
    </file>
    <filePlan>
        <code>54321</code>
        <recordCategoryCode>12</recordCategoryCode>
        <recordCategoryCode>14</recordCategoryCode>
        <dispositionAutority>String</dispositionAutority>
        <userDefinedData>
            <dataFieldName>String</dataFieldName>
            <dataFieldValue>String</dataFieldValue>
        </userDefinedData>
        <userDefinedData>
            <dataFieldName>String</dataFieldName>
            <dataFieldValue>String</dataFieldValue>
        </userDefinedData>
    </filePlan>
    <recordCategory>
        <code>12</code>
        <name>Financial</name>
        <description>Top secret documents</description>
        <dispositionInstructionCode>20</dispositionInstructionCode>
        <userDefinedData>
            <dataFieldName>Security Classification</dataFieldName>
            <dataFieldValue>1A</dataFieldValue>
        </userDefinedData>
    </recordCategory>
    <recordCategory>
        <code>14</code>
        <name>Minutes</name>
        <description>Minutes</description>
        <dispositionInstructionCode>22</dispositionInstructionCode>
    </recordCategory>
    <dispositionInstruction>
        <name>String</name>
        <code>20</code>
        <type>Time</type>
        <actionType>Transfer</actionType>
        <actionDescription>Transfer to the National Archives of the United States </actionDescription>
        <retentionPeriod unit="month">3</retentionPeriod>
    </dispositionInstruction>
    <dispositionInstruction>
        <name>String</name>
```

```
        <code>22</code>
        <type>Event</type>
        <eventDescription>Case completion</eventDescription>
        <actionType>Destroy</actionType>
        <retentionPeriod unit="year">3</retentionPeriod>
    </dispositionInstruction>
    <fileCategory>
        <code>101</code>
        <name>Financial</name>
        <description>Financial</description>
        <cutoffInstructionCode>333</cutoffInstructionCode>
    </fileCategory>
    <cutoffInstruction>
        <code>333</code>
        <comment>Cutoff the inactive file at the end of each fiscal (or calendar) year; then apply the retention
period.</comment>
        <periodic unit="year">1</periodic>
    </cutoffInstruction>
</RMASystem>
```

6 Bibliography

[1] Brickley and Guha, “Resource Description Framework (RDF) Schema Specification 1.0,” W3C Candidate Recommendation, March, 2000. http://www.w3.org/TR/rdf-schema/ (last accessed 3/12/2002).

[2] Cummings. “Records Management: Management Now and in the Future,” June 28, 2001. http://www.nara.gov/era/cendi/cendi_index.html (last accessed 3/12/2002).

[3] Department of Defense 5015.2-STD, “Design Criteria Standard for Electronic Records Management Software Applications,” November, 1997.

[4] Duckett, et al. “Professional XML Schema,” Wrox Press, Ltd., July, 2001.

[5] http://www.nara.gov/records/ (last accessed 3/12/2002).

[6] http://www.w3c.org/XML/ (last accessed 3/12/2002).

[7] International Electrotechnical Commission 62045, “Management Data (Metadata) for Technical Documents, Part 1,” July, 1999.

[8] Prescott, Underwood and Kindl. “Functional Baseline Requirements and Data Elements for Records Management Application Software,” Army Research Laboratory, Software Technology Branch, August, 1995.

[9] PROS 99/007 Specification 2, “VERS Metadata Scheme Version 1.0,” Public Record Office Victoria, April, 2000.

[10] Section 1234.22(b) of title 36, Code of Federal Regulations, ”Creation and Use of Text Documents”

[11] Section 3301 of title 44, United States Code, “Definition of Records”

[12] Thibodeau. “Building the Electronic Records Management Information Infrastructure,” April 2001. http://www.nara.gov/era/arma/arma_index.html (last accessed 3/12/2002).

[13] Whitehead. “Workshop on Metadata Interoperability for Electronic Records Management,” notes, slides and other materials from http://www.soe.ucsc.edu/~ejw/metadata, November 15, 2001. (URL last accessed 3/12/2002).

[14] XML Spy v4.3. Available from http://www.xmlspy.com/ (last accessed 3/11/2002).

7 Appendix

This table shows a draft mapping between the metadata elements in VERS and those in DoD 5015.2. While this is incomplete, it should provide a sense of how the schema defined in this document could form the core of an implementation for a standard other than DoD 5015.2.

Victorian Electronic Records Strategy (VERS)		Maps to	
M37	Subject	R1	subject
M54	Date/Time Registered	R2	fileDate
M107	Recipient	R3	transmissionReceiptData
M84	Current Location	R4	location
M62	Data Format	R5	format
M63	Medium	R6	media
M55	Date/Time Created	R7	creationDate
M106	Originator	R8	originator
M12	Agent	R9	originatingOrg
M92	Disposal Status (assigned value = “Permanent”)	R10	vital
M102	File Identifier	R11	fileCode
M88	Disposal	R12	recordCategoryCode
M103	Vers record Identifier	R13	ID
M42	M43 Related Item ID	R15	relatedRecords
M42	M43 Related Item ID, M44 Relation Type (assigned value = “Next”)	R16	sucessor
M42	M43 Related Item ID, M44 Relation Type (assigned value = “Previous”)	R16	predecessor
M91	Disposal Action Due	R17	dispositionActionDate
M90	Sentence	DI3,DI4,DI4	type, actionType, retentionPeriod
M88	Disposal	F1	filePlanCode
M103	Vers record Identifier	F2	code
M42	M43 Related Item ID, M44 Relation Type (assigned value = “Contained in”)	F3	
M144	Date/Time Closed	F5	closureDate
M90*	Sentence	File Plan	
M88*	Disposal	Record Category	
M144*	Date/Time Closed	File Category, Cutoff Instruction	

*These items are similar to items in DoD 5015.2, but don’t directly map.