MoReq2: A European Contribution to the Preservation of Electronic Records

Marc Fresko
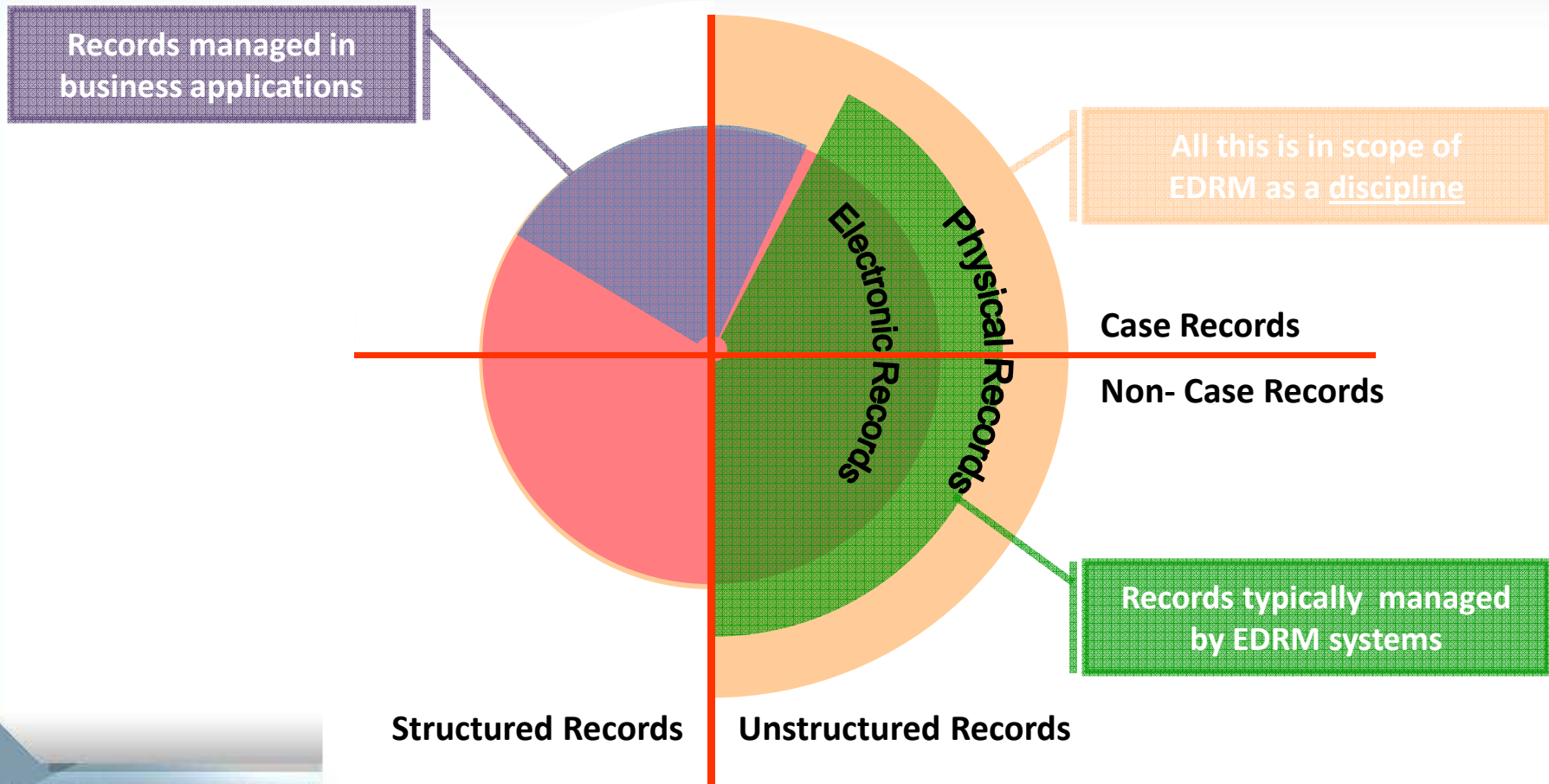Inforesight Limited
marc.fresko@inforesight.co.uk
+44 20 8645 0080

Agenda

- Scope: electronic records management systems
- MoReq2
- Recognition of "components"
- Migration
- Automated rendition
- Import and export features
- Preservation metadata
- XML schema

Kinds of electronic records
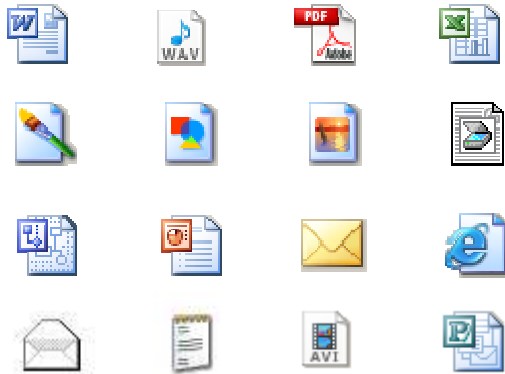
- There are many kinds of electronic records
- In principle, any of them could be managed using an electronic document and records management system (EDRMS)
- In practice, not all of them are

Kinds of electronic records

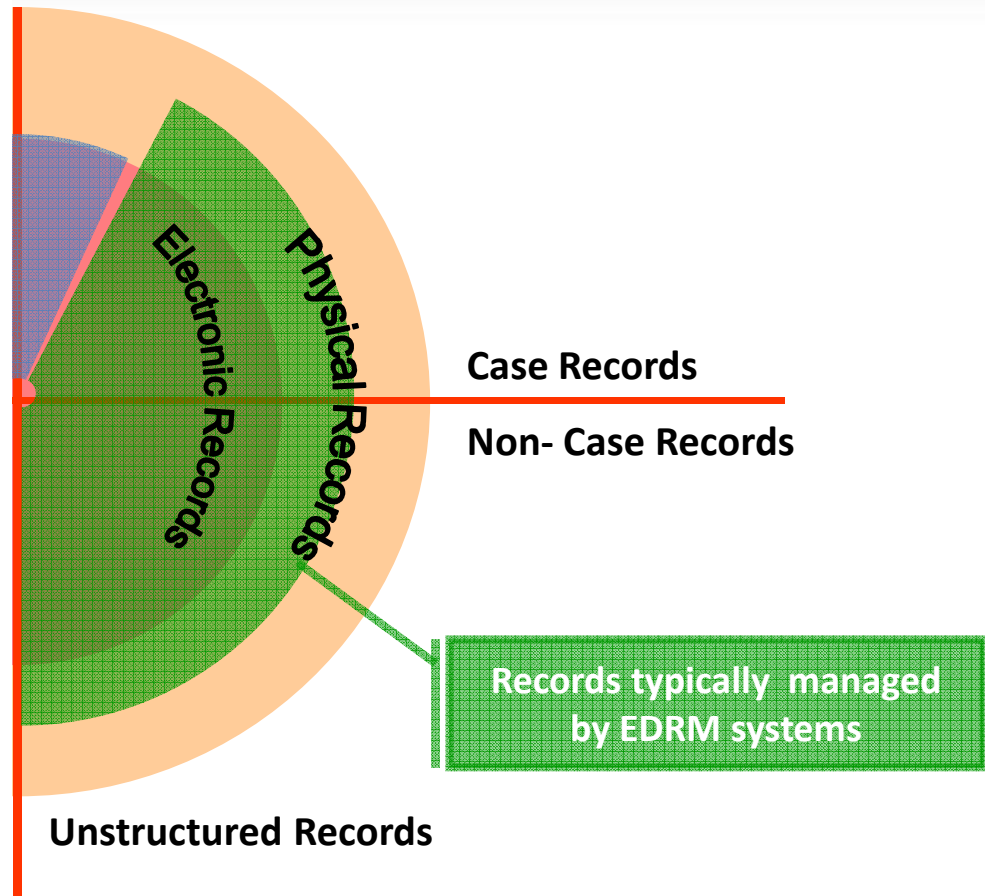Records managed in business applications

All this is in scope of EDRM as a discipline

Electronic Records

Physical Records

Case Records

Non- Case Records

Records typically managed by EDRM systems

Structured Records   Unstructured Records

inforesight

Records typically managed using EDRMSs

▲ Unstructured

▲ Mainly non-case records

▲ Some case records

Electronic Records

Physical Records

Case Records

Non- Case Records

Records typically managed by EDRM systems

Unstructured Records

inforesight

5

Agenda

- Scope: electronic records management systems
- MoReq2
- Recognition of "components"
- Migration
- Automated rendition
- Import and export features
- Preservation metadata
- XML schema

inforesight

MoReq2, MoReq2

- MoReq =
  Model **Req**uirements for the management of electronic records

- Conceptually similar to US DoD 5015.2
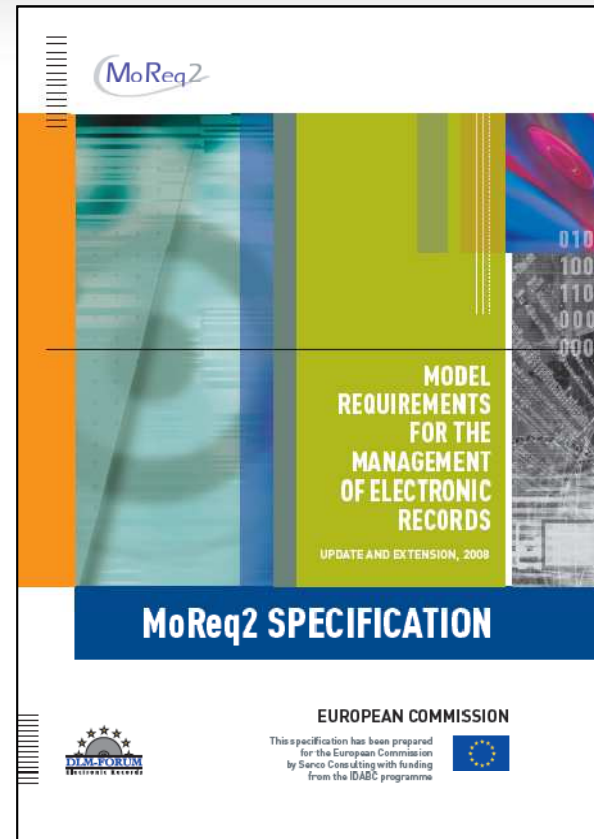 (but "conceptually" is the only sort of similarity)

MoReq

- March 2001
- Invented by DLM Forum
- Funded by EC
- Developed by Cornwell
- Translated by many
- Orphaned by all

DLM Forum

- Redesigned (and re-invigorated) after MoReq
- Conceived and launched MoReq2 project in 2006
- MoReq2 published in early 2008

MoReq2

MODEL
REQUIREMENTS
FOR THE
MANAGEMENT
OF ELECTRONIC
RECORDS

UPDATE AND EXTENSION, 2008

MoReq2 SPECIFICATION

EUROPEAN COMMISSION

This specification has been prepared
for the European Commission
by Serco Consulting with funding
from the IDABC programme

DLM-FORUM
Electronic Records

MoReq2 project was supported by:

Adobe · Anite · BCC SERVICES · Captaris · DEXMAR · DocuWare

EMC² · e SMART systems · Fabasoft · T··Systems· · HAESSLER

HYPERWAVE · IBM · EXPRIT · ICZ · IDOX · MELONWEB · meridio

Microsoft · Objective · OPEN TEXT CORPORATION · ORACLE · SAP

SER · SPESCOM · VIGNETTE · XEROX · ca · TOWER SOFTWARE

Also later joiners: Capgemini, Civica, Fujitsu, EDRM Solutions, ErgoGroup, FreeSoft, Getronics, Lockheed Martin, Morse…

inforesight

10

MoReq2 is influential in and beyond Europe

- MoReq was used around the world, incl. i1n USA
- Many MoReq2 translations started
- Conferences, seminars, from Bogotá to Moscow
- Interest outside Europe – for example
  - Russian translation published
  - Korean translation nearly complete
  - Worldwide press coverage

inforesight

欧盟《电子文件管理通用需求第二版——MoReq2》解析

□张宁

　　随着世界信息化程度的加深,电子文件管理日渐成熟。从世界范围看,在 ISO15489 成为文档管理的第一个国际标准之后,ISO23081、ISO22310、ISO26122、ISO26012 等国际标准也相继出台或正在制定中。这些国际标准的制定,表明了文档管理,尤其是电子文件的管理,正逐渐步入规范化、标准化的发展轨道。

　　在国际大趋势的影响下,鉴于文档管理的个性化特征,目前若干国家和地区纷纷开始制定本国或本地区的电子文件管理标准,以期解决文件管理中一些操作层面共性与个性相互协调的问题,电子文件管理系统功能需求的设计便是其中最为关键的一环。除了大家熟知的美国 ERA 项目外,欧盟的《电子文件管理通用需求》(Model Requirements for the Management of Electronic Records,以下简称 MoReq)也很有代表性。2008 年 2 月该需求出台

　　为了使 MoReq 适应欧盟各国新阶段文件管理的需求,更加贴近于现阶段和未来一段时期电子文件管理的实践,2006 年 2 月,欧盟正式启动 MoReq 的修订程序,并发布了 MoReq2《范围报告》(Scoping Report)。

　　该报告指出,MoReq2 的目标是适用于每一个欧盟成员国,这就要求它所提供的模型具有一定的灵活性,能够在不同的环境中均能够保证电子文件质量。根据此目标,MopReq2 内容的重点将在于:(1)制定可信的电子文件管理基准模型 (The base module constitutes the minimum necessary to provide credible electronic records management),为不同环境下的应用系统提供基本需求框架,包括利用控制、文件保管期限和处置、输出和移交、保管及元数据。(2)建立集成化的信息系统。仅限于关注电子管理系统是不够的,得多系统(如内容管理系统、知识管理系统、文

MoReq2 and DoD 5015.2

▲ Difficult to compare

▲ Any detailed reconciliation will be highly subjective (and so of limited value)

▲ MoReq2 has much broader functional scope

▲ There are incompatible requirements (including for preservation)

PRO	MoReq	5015.2
PRO: Revised requirements	**MoReq : 2001**	**DOD 5015.2 : 2002**
Record Organisation		
Classification scheme / fileplan		
A.1.1	(3.1.1)	
A.1.2	3.1.2, (3.1.3)	
A.1.3	3.1.5	
A.1.4	3.1.6	
A.1.5	3.2.9	
A.1.6	3.4.1	
** A.1.7	-	
** A.1.8	-	
A.1.9	-	
** A.1.10	э 3.1.8	
* A.1.11	э 3.1.9	
Class metadata		
A.1.12	(3.2.1)	
A.1.13	-	
A.1.14	3.2.2	
A.1.15	(3.2.2)	
A.1.16	(3.1.4)	
A.1.17	-	
A.1.18	-	
** A.1.19	-	
A.1.20	3.2.5	
A.1.21	э 12.1.11	
** A.1.22		

MoReq2 includes "basic" preservation features

- Physical environment
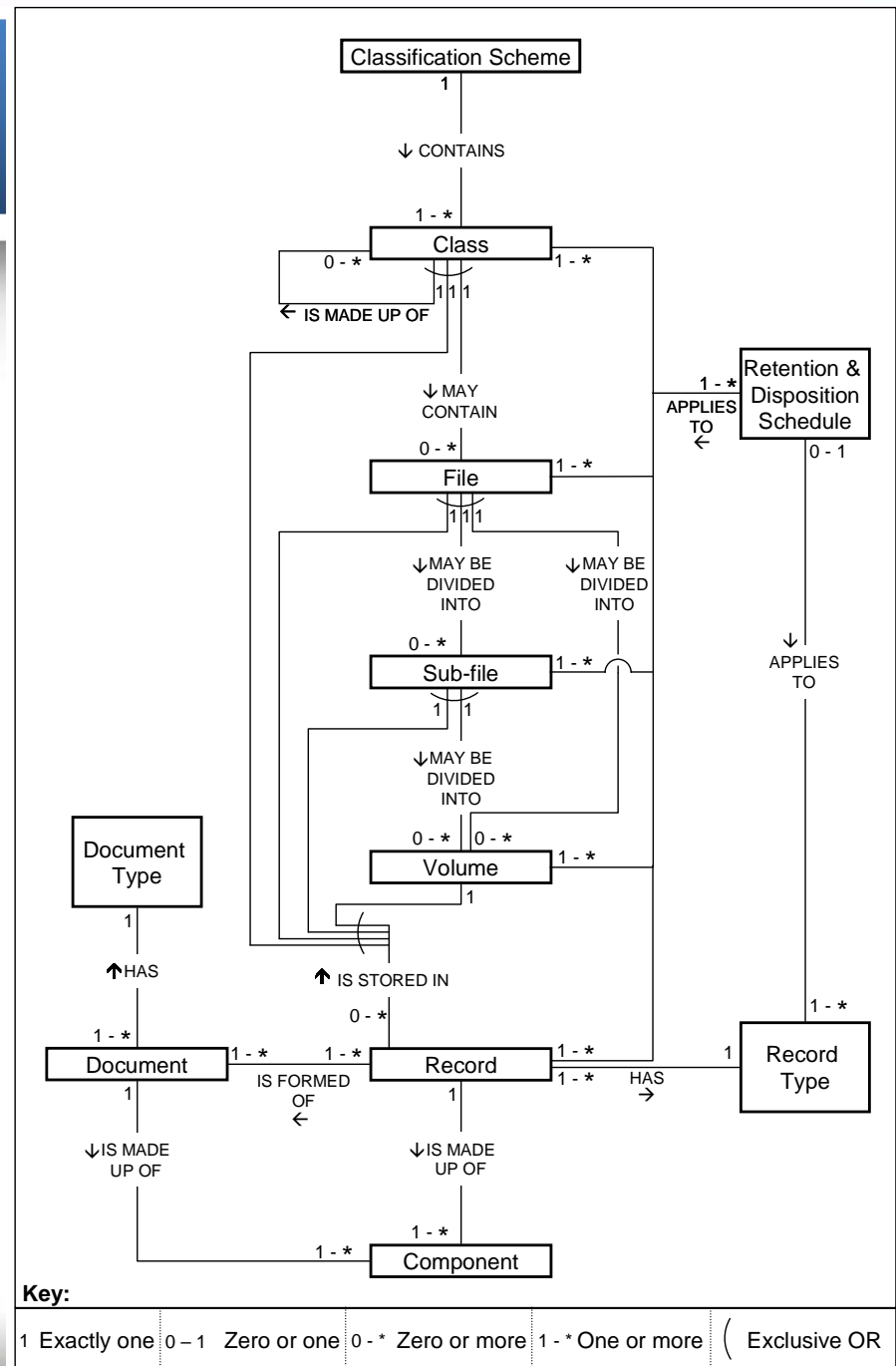- Monitoring of error rates
- Comparison of copies

inforesight

Agenda

- Scope: electronic records management systems
- MoReq2
- Recognition of "components"
- Migration
- Automated rendition
- Import and export features
- Preservation metadata
- XML schema

Components
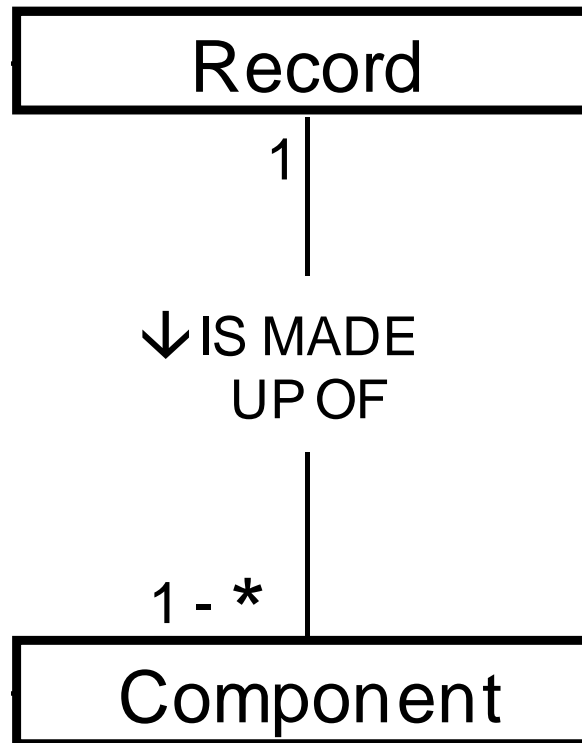
- Individual "files" (in the IT sense not the RM sense) that make up a record
- Well understood in preservation work
- Not well understood in other records management specifications

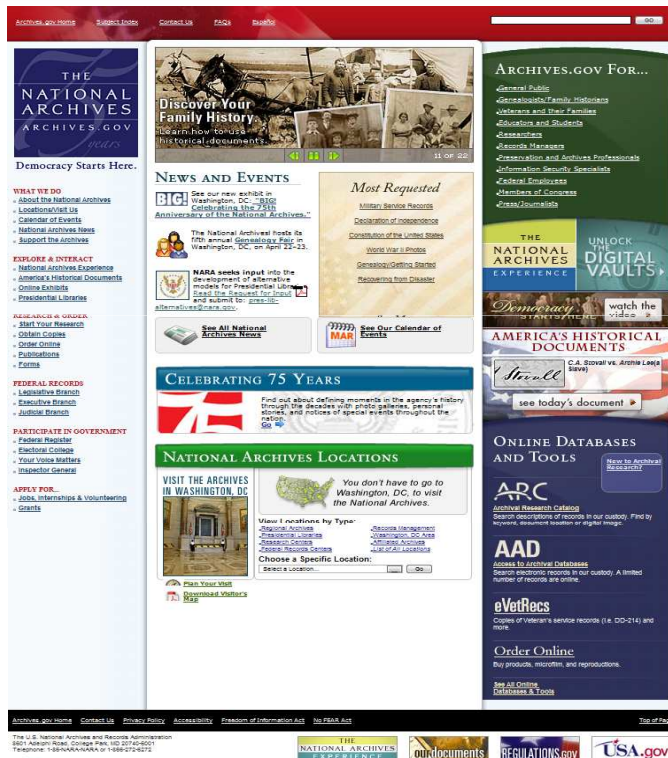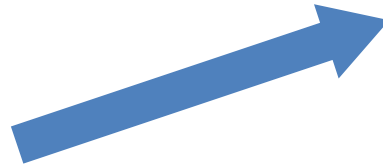# MoReq2 entity-relationship model

Classification Scheme
1

↓ CONTAINS

1 - *
Class
0 - *   1|1|1   1 - *
← IS MADE UP OF

↓ MAY CONTAIN

1 - *   Retention &
APPLIES   Disposition
TO   Schedule
←   0 - 1

0 - *
File
1 - *
1|1|1

↓ MAY BE DIVIDED INTO   ↓ MAY BE DIVIDED INTO

0 - *
Sub-file
1 - *
1|1

↓ MAY BE DIVIDED INTO

APPLIES TO

0 - * 0 - *
Volume   1 - *
1

↑ HAS   ↑ IS STORED IN

Document Type
1

0 - *

1 - *
Document   1 - *   1 - *
1   IS FORMED OF
            ←

Record   1 - *
           1 - *   HAS
1   →

1 - *
Record Type

↓ IS MADE UP OF   ↓ IS MADE UP OF

1 - *
1 - *   Component

Key:

1 Exactly one	0 – 1 Zero or one	0 - * Zero or more	1 - * One or more	(Exclusive OR

Components

Record

1

↓ IS MADE UP OF

1 - *

Component

Components: example

NARA web page

IS MADE UP OF

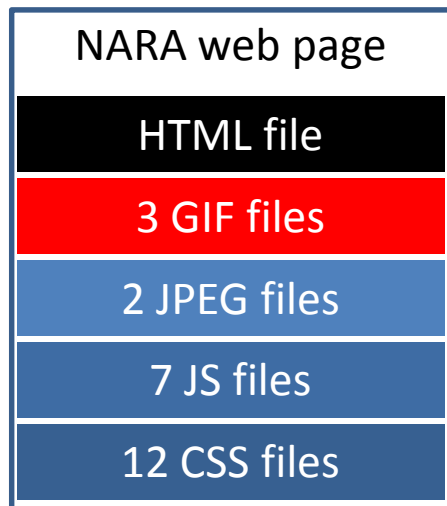| |
|---|
| 1 HTML file |
| 3 GIF files |
| 2 JPEG files |
| 7 JS files |
| 12 CSS files |

inforesight

Agenda

▲ Scope: electronic records management systems

▲ MoReq2

▲ Recognition of "components"

▲ Migration

▲ Automated rendition

▲ Import and export features

▲ Preservation metadata
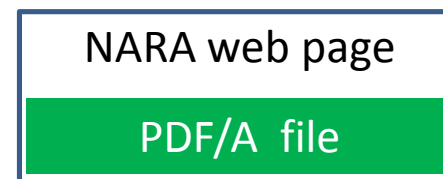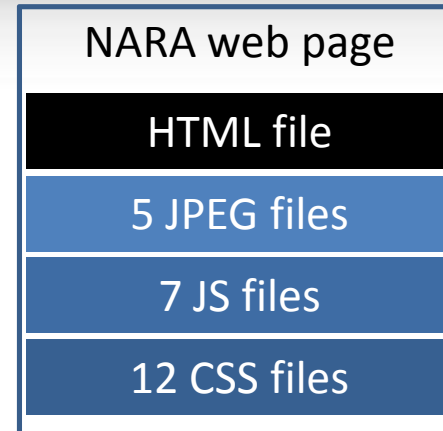
▲ XML schema

inforesight

Migration (aka "rendition")

- Current orthodoxy holds that migration of information from existing formats to preferred preservation formats may play an important part in active preservation measures
- Possible examples:
  - .wp5 to PDF/A
  - .odg to JPEG

inforesight

Complex relationships

▲ Example:

NARA web page

HTML file
3 GIF files
2 JPEG files
7 JS files
12 CSS files

or

NARA web page

HTML file
5 JPEG files
7 JS files
12 CSS files

NARA web page

PDF/A file

MoReq2 migration features

- Report on files formats
- Migrate records in selected file formats
- Migrate components in selected file formats
- Capture of reason for migration
- Retention of pre- and post-migration formats
- OAIS DIP compatibility
- Maintenance of record integrity
  - Metadata
  - Links

Record integrity

- Migration of compound records can result in the need to change record contents
- For example in a web page
  - Change
  - To
- This is slightly controversial

Agenda

- Scope: electronic records management systems
- MoReq2
- Recognition of "components"
- Migration
- Automated rendition
- Import and export features
- Preservation metadata
- XML schema

Automated rendition

- Requires that system be able to migrate records upon capture to a specified preservation file format
- A natural extension of the above
- Allows a Records Manager or other Information Manager to build a collection of records that are preservable from the outset

Agenda

- Scope: electronic records management systems
- MoReq2
- Recognition of "components"
- Migration
- Automated rendition
- Import and export features
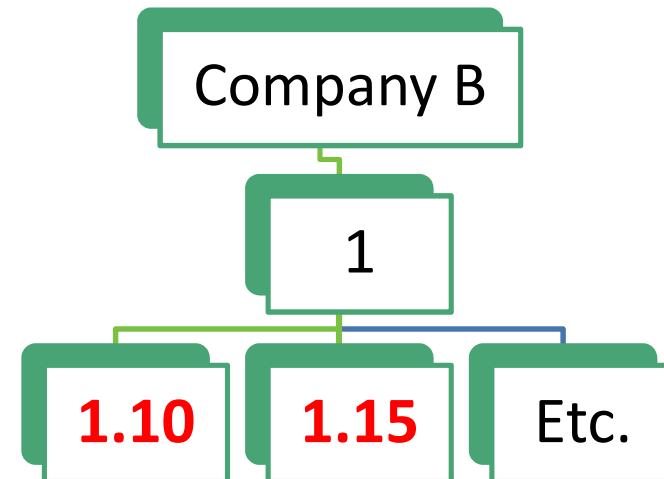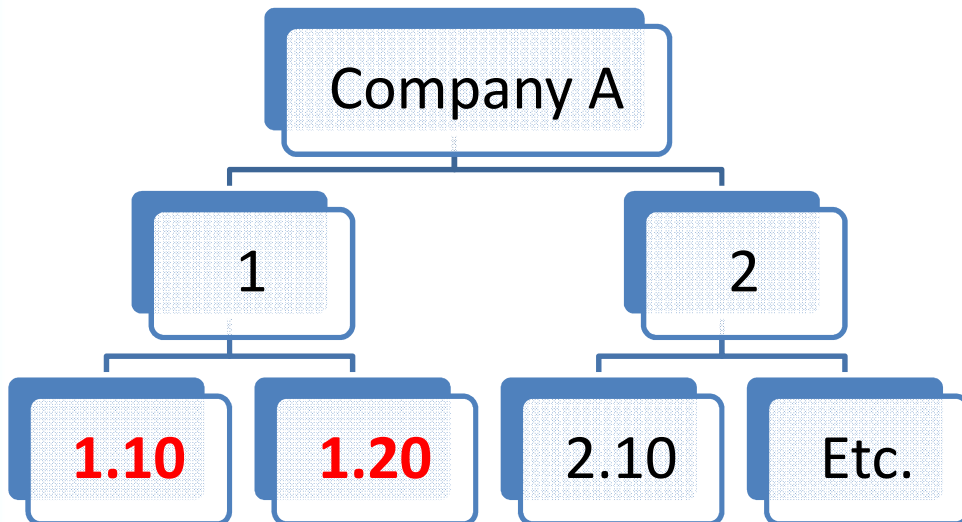- Preservation metadata
- XML schema

inforesight

The need for import/export features

- Where digital preservation matters, it is almost certain that preservation period will exceed the lifetime of the EDRM system

- Therefore the EDRM system will need to export the records

  - to another EDRM system, or

  - to a digital preservation system

- MoReq2 includes requirements for such export

Import/export features

- Deals with not just records, but also the hierarchy of aggregations
- Requirements for protocol to notify success or failure of a transfer
- Requirements to deal with incomplete or inconsistent metadata
- Requirements to deal with naming and numbering clashes

One example (from many) of a conflict

Company A

```
Company A
├── 1
│   ├── 1.10
│   └── 1.20
└── 2
    ├── 2.10
    └── Etc.
```

Company B

```
Company B
└── 1
    ├── 1.10
    ├── 1.15
    └── Etc.
```

inforesight

Agenda

- Scope: electronic records management systems
- MoReq2
- Recognition of "components"
- Migration
- Automated rendition
- Import and export features
- Preservation metadata
- XML schema

inforesight

Metadata model

- MoReq2 includes a comprehensive metadata model
- Tightly defined to support interoperability (as opposed to 5015.2, which is not)
- ISO 23081-compliant
- Metadata defined for records and all aggregations
- Also for retention and disposition schedules, disposal holds, users, user groups and roles
- Not tested
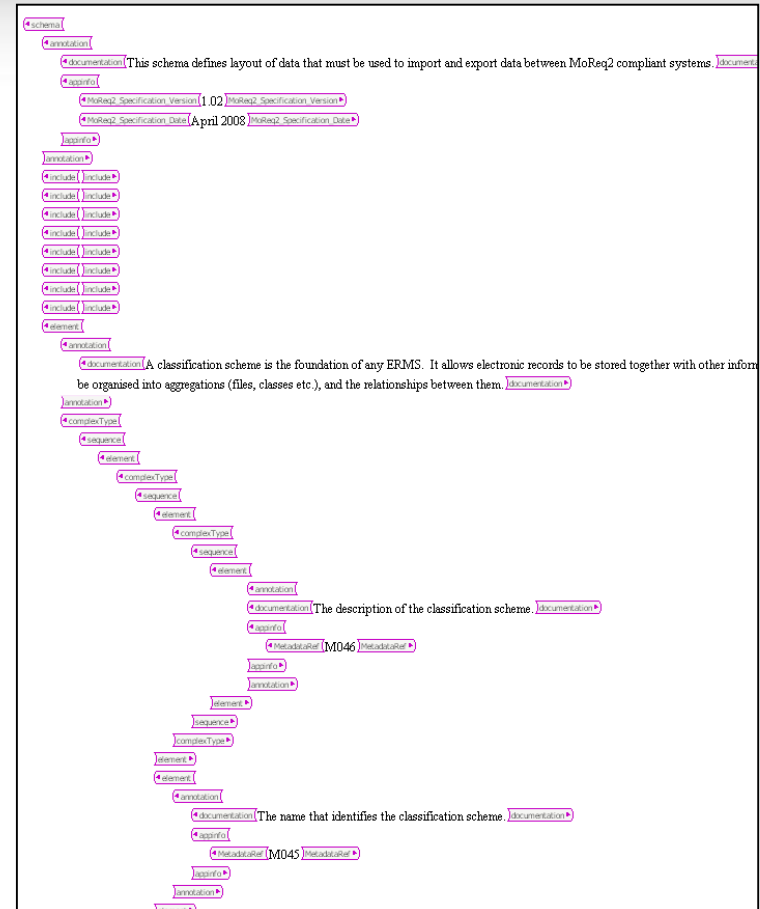- May contain errors

inforesight

32

Preservation metadata

▲ Preservation-related metadata elements for components:

- ▼ M131: Description.abstract.reason_for_rendition
- ▼ M130: Identity.system_identifier_rendition
- ▼ M150: Relation.has_rendition
- ▼ M151: Relation.is_rendition_of
- ▼ M128: Use.technical_environment.file_format
- ▼ M133: Use.technical_environment.file_format_original
- ▼ M129: Use.technical_environment.file_format_version
- ▼ M142: Use.technical_environment.file_format_version_original

Agenda

- Scope: electronic records management systems
- MoReq2
- Recognition of "components"
- Migration
- Automated rendition
- Import and export features
- Preservation metadata
- XML schema

XML schema

▲ Supports import and export

▲ Includes records and aggregations (files, classes etc.)

▲ Not tested

Conclusion

▲ MoReq2 is a standard specification for an EDRM office application

▲ It is already, and will continue to be, very influential

▲ MoReq2 does not define a preservation system

▲ MoReq2 adds preservation features to EDRM

▲ Adoption of preservation features in other "front-end" systems is surely an important step for the future

Questions?

Marc Fresko
Inforesight Limited

✉ marc.fresko@inforesight.co.uk

☎ +44 20 8645 0080

🌍 http://www.inforesight.co.uk

inforesight