

# Livre Blanc. 4



Standards Council of Canada  
Conseil canadien des normes



*afnor*  
NORMALISATION



**ISO 30300 - 30301 : métadonnées pour les documents d'activité. Principes directeurs, vue conceptuelle et modalités**

Juin 2013

**ISO 30300 - 30301 : Metadata for records, guiding principles, conceptual overview and modalities**

June 2013

## Livre blanc 4

# ISO 30300 – 30301 : Métadonnées pour les documents d'activité, principes directeurs, vue conceptuelle et modalités

## ISO 30300 – 30301 : Metadata for records, guiding principles, conceptual overview and modalities

### Sommaire / contents

<b><u>1</u></b>	<b><u>AVANT-PROPOS : POURQUOI CE LIVRE BLANC</u></b>	<b><u>3</u></b>
<b><u>2</u></b>	<b><u>LE CONTEXTE NORMATIF ACTUEL</u></b>	<b><u>6</u></b>
<b><u>3</u></b>	<b><u>METADONNEES RAPPELS ET NORMES EXISTANTES</u></b>	<b><u>7</u></b>
3.1.	DEFINITIONS, RAPPELS DES FONCTIONS ET DES SPECIFICITÉS DES MÉTADONNÉES	7
3.2.	LES NORMES EXISTANTES, LEUR LIMITES ET LEUR PERTINENCE	12
<b><u>4</u></b>	<b><u>REVERSIBILITÉ, SÉCURITÉ, CLOUD : QUELLES EXIGENCES IMPACTANT LES MÉTADONNÉES ?</u></b>	<b><u>15</u></b>
4.1.	EXIGENCE DE RÉVERSIBILITÉ	15
4.2.	EXIGENCE DE SÉCURITÉ	15
4.3.	COMMENTAIRES VIS-À-VIS DU CLOUD STORAGE	17
<b><u>5</u></b>	<b><u>ÉTUDES DE CAS ET POINTS DE VUE</u></b>	<b><u>17</u></b>
5.1.	LINKLATERS, CABINET D'AVOCAT INTERNATIONAL(SYLVIE FORASTIER - LUXEMBOURG) ...	17
5.2.	POINT DE VUE SUR LES MÉTADONNÉES EN ENTREPRISE (LUCAS COLET - LUXEMBOURG) .....	19
5.3.	LES MÉTADONNÉES ET LEUR RÔLE DANS LA GESTION ÉLECTRONIQUE DES DOCUMENTS – UN APERÇU DE L'EXPÉRIENCE DE LA COMMISSION EUROPÉENNE (MATTIAS HALLIN) .....	20
5.4.	LA MISE EN OEUVRE DE ISO 23081 DANS LE CONTEXTE DU GOUVERNEMENT DU CANADA (NATHALIE VILLENEUVE - CANADA) .....	24
5.5.	POINT DE VUE DE SUR LE CLOUD STORAGE (JOHANNE. LESSARD – QUEBEC) .....	25
<b><u>6</u></b>	<b><u>CONCLUSION</u></b>	<b><u>28</u></b>

<b><u>8</u></b>	<b><u>ENGLISH VERSION .....</u></b>	<b><u>29</u></b>
<b><u>9</u></b>	<b><u>FOREWORD: THE REASON FOR THIS WHITE PAPER.....</u></b>	<b><u>29</u></b>
<b><u>10</u></b>	<b><u>THE CURRENT STANDARDS CONTEXT .....</u></b>	<b><u>32</u></b>
<b><u>11</u></b>	<b><u>METADATA, REMINDERS AND EXISTING STANDARDS .....</u></b>	<b><u>33</u></b>
11.1.	DEFINITIONS, REMINDERS OF FUNCTIONS AND SPECIFIC CHARACTERISTICS OF METADATA	33
11.2.	EXISTING STANDARDS, THEIR LIMITATIONS AND THEIR RELEVANCE .....	37
<b><u>12</u></b>	<b><u>REVERSIBILITY, SECURITY, CLOUD: WHICH REQUIREMENTS IMPACT METADATA? .....</u></b>	<b><u>40</u></b>
12.1.	REVERSIBILITY REQUIREMENT.....	40
12.2.	SECURITY REQUIREMENT .....	41
12.3.	COMMENTS REGARDING CLOUD STORAGE.....	42
<b><u>13</u></b>	<b><u>CASE STUDIES AND VIEWPOINTS.....</u></b>	<b><u>42</u></b>
13.1.	LINKLATERS, INTERNATIONAL LAW FIRM (SYLVIE FORASTIER - LUXEMBOURG).....	42
13.2.	VIEWPOINT ON METADATA IN ENTERPRISES (LUCAS COLET - LUXEMBOURG).....	45
13.3.	METADATA AND THEIR ROLE IN ELECTRONIC DOCUMENT MANAGEMENT – AN OVERVIEW OF THE EXPERIENCE AT THE EUROPEAN COMMISSION (MATTIAS HALLIN) .....	46
13.1.	THE IMPLEMENTATION OF ISO 23081 IN THE GOVERNMENT OF CANADA (GC) CONTEXT (NATHALIE VILLENEUVE) .....	49
13.2.	VIEWPOINT ON CLOUD STORAGE (JOHANNE. LESSARD – QUEBEC).....	51
<b><u>14</u></b>	<b><u>CONCLUSION.....</u></b>	<b><u>53</u></b>
<b><u>15</u></b>	<b><u>LE GROUPE DE TRAVAIL – WORKING GROUP .....</u></b>	<b><u>54</u></b>

## 1 AVANT-PROPOS : POURQUOI CE LIVRE BLANC

---

Un travail de normalisation de grande ampleur est actuellement en cours concernant la gestion des documents d'activité. Une série de Livres blancs publiés par la communauté francophone a eu pour objet de contribuer à ce travail.

**Tout d'abord le Livre blanc n°1.** Il a été présenté à la réunion du TC46/SC11 qui s'est tenue à Sidney en 2011 et a permis d'introduire la série des normes de système de management ISO 3030X. Il a permis également d'exposer les nécessaires évolutions de l'ISO 15489 du *records management*. Ce premier livre blanc a permis le départ d'une dynamique francophone matérialisée par la version en français de l'ISO 30300 et 30301 qui a été votée à l'ISO. Il a permis aussi de faire connaître et d'affirmer le rôle de la CN11 comme force de proposition.

**Ensuite, le Livre blanc n°2.** Il a été présenté à la réunion du TC46/SC11 qui s'est tenue à Londres en novembre 2011. Son influence a aidé tout particulièrement à la mise en révision de l'ISO 15489. L'équipe de normalisation du Canada a rédigé une synthèse en anglais.

**Enfin le Livre blanc n°3, bilingue.** Ce Livre blanc avait pour thème les modélisations et diagrammes représentant les systèmes de gestion de documents d'activité. Il a été présenté à la réunion du TC46/SC11 qui s'est tenue à Berlin en mai 2012. Cette version, entièrement traduite en anglais, a été l'objet de nombreuses contributions internationales. Parallèlement, en France elle a permis de lancer plusieurs réflexions, et en particulier sur la pertinence du paradigme des trois âges en matière d'archivage eu égard à la représentation numérique des documents d'activité.

**Le présent Livre blanc n°4, bilingue** a pour ambition d'aborder la question des métadonnées et de leur normalisation afin de préparer la réunion du TC46/SC11 qui se tiendra à Paris en juin 2013. Il s'agit d'un domaine particulièrement sensible puisque les « données sur les données » permettent d'identifier, de retrouver et d'échanger de l'information entre les systèmes. De plus, certaines métadonnées peuvent contribuer à donner une valeur de preuve. De fait et avec l'arrivée du *Big Data* et de l'activité économique autour du *Cloud Computing*, la gestion et préalablement la définition des métadonnées ont de fait un caractère stratégique.

La norme ISO 23081 : Métadonnées pour les *records* de janvier 2006 définit les principes qui régissent les métadonnées liées aux documents d'activité et à leur gestion. Dans son introduction la norme précise : « L'ISO 23081 établit un cadre permettant de créer, gérer et utiliser les métadonnées associées à la gestion des documents d'activité et explique les principes qui les régissent ».

Le processus de révision systématique de l'ISO 23081 intervenant en 2014, comme celui de la révision de l'ISO 15489 commencé en 2012, constituent une opportunité exceptionnelle pour conduire de nouvelles réflexions.

C'est dans ce contexte que ce Livre blanc souhaite trouver sa place en essayant d'apporter des éléments de réponse à un certain nombre de questionnements.

## IL Y A-T-IL UNE SPÉCIFICITÉ DES MÉTADONNÉES LIÉES À LA GESTION DES DOCUMENTS D'ACTIVITÉ ?

---

En matière de métadonnées et plus largement de mécanismes de repérage d'informations et de documents dans des systèmes, de nombreuses normes existent déjà. L'ISO 15836 dite du *Dublin Core* ou celle d'échange d'information comme la norme Z39.50 ou encore l'ISO 23950 ne peuvent être ignorées. La question qui se pose est de savoir si la gestion des documents d'activité nécessite des métadonnées spécifiques : en d'autres termes, un jeu de métadonnées sans lequel la série de normes ISO 30300 et les normes techniques associées particulièrement l'ISO 23081 et l'ISO 15489 ne pourraient être mise en œuvre de façon fiable, sécurisée et pérenne.

Ce Livre blanc abordera également les réflexions actuellement en cours relatives à la création d'un identifiant unique, sorte d'ISBN<sup>1</sup> des documents d'activité ou d'indicateur persistant (cf. DIN 31646).

Dans ce contexte d'évolution du corpus normatif de gestion des documents d'activité, deux principales hypothèses de travail se présentent afin de faire évoluer les questionnements vis-à-vis des métadonnées:

- Réviser la norme ISO 23081 afin d'obtenir une définition plus précise du jeu de métadonnées pour finalement permettre une mise en œuvre plus aisée de la norme,
- Intégrer dans la révision de la norme ISO 15489 des éléments sur les exigences relatives au jeu de métadonnées, comme par exemple créer un volet 3 spécifique aux métadonnées.

Ce Livre blanc tentera d'apporter de nouveaux éclairages à ces questions et de formaliser un point de vue.

## COMMENT ENVISAGER L'ÉVOLUTION DES MÉTADONNÉES, EU ÉGARD AU CYCLE DE VIE ?

---

Les documents d'activité s'inscrivent dans un cycle de vie au cours duquel interviennent des événements métiers auxquels ils sont liés ainsi que des évolutions technologiques qui s'imposent, comme des changements de support physique par exemple. A l'occasion de ces événements de gestion, il est nécessaire de faire évoluer les métadonnées sans remettre en cause l'intégrité du document et la réversibilité du processus. De fait, le couple « document d'activité » devient au cours du temps et des changements successifs de support un système d'emboitage de type poupées russes qui peut s'écrire (((document-métadonnées)-métadonnées)-métadonnées), etc.

La norme ISO 23081 formule ainsi cette nécessité « Toutes les métadonnées liées au document d'activité ainsi que celles qui s'ajoutent au cours de sa gestion et de son utilisation, forment aussi un document d'activité, document des métadonnées qu'il importe aussi de gérer ».

De fait, le couple cycle de vie/métadonnées est bien spécifique à la gestion des documents d'activité et sera également commenté dans ce Livre blanc.

---

<sup>1</sup>International Standard Book Number.

Enfin, depuis quelque temps l'expression *Cloud Computing*<sup>2</sup> est apparue à grand renfort de démarches commerciales et d'offres de services de la part de fournisseurs. Il dispose de différentes déclinaisons dont le *Cloud Storage* qui ne peut être ignoré en matière de gestion des documents d'activité. Sur ce sujet, les membres du groupe de travail ont souhaité faire quelques commentaires.

Ce Livre blanc comprend trois principaux chapitres qui analysent et tentent de répondre aux questionnements précédents :

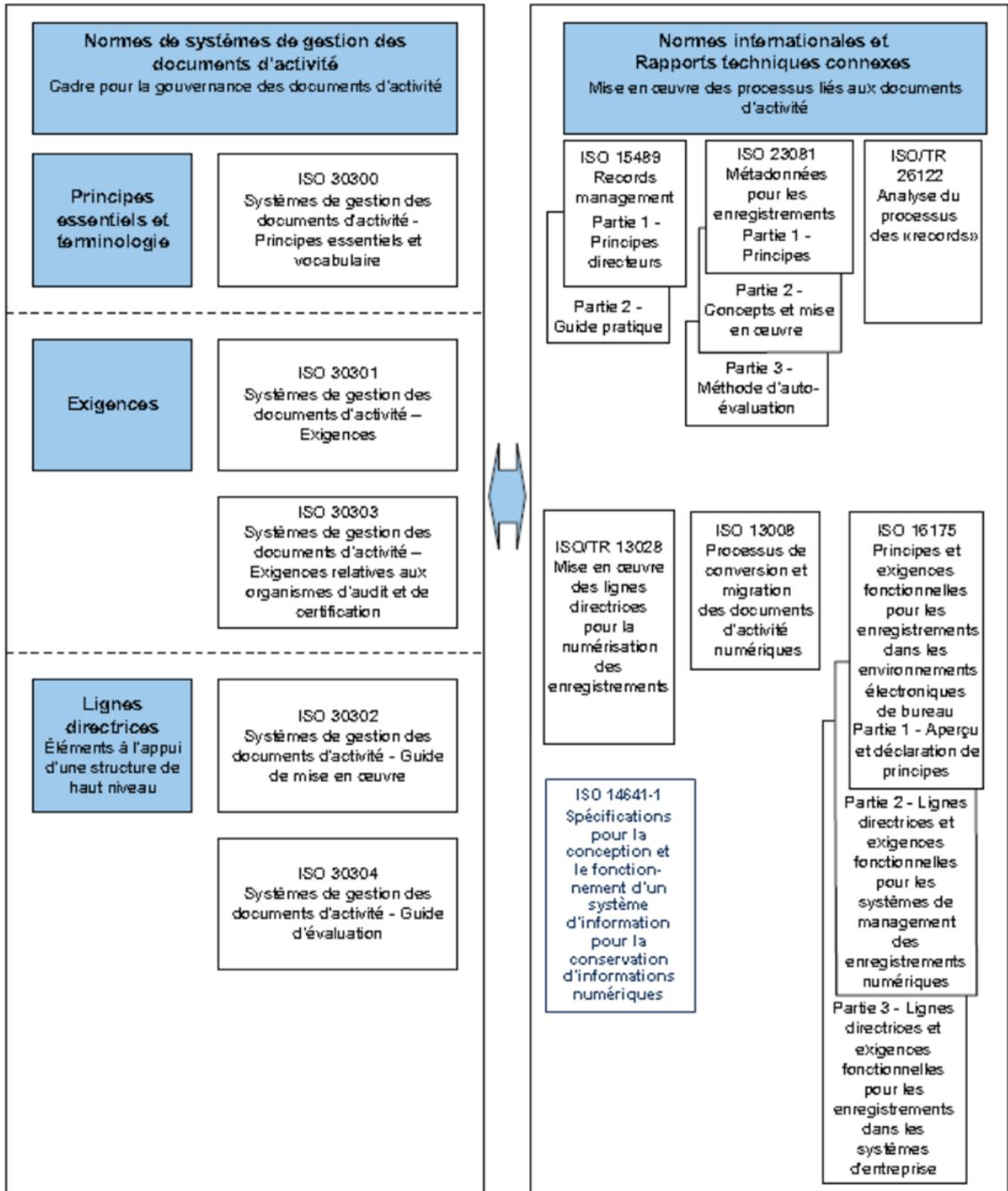
- Le chapitre 3 réunit des définitions et rappelle ce que sont des métadonnées.
- Le chapitre 4 fait état d'exigences concernant des métadonnées de gestion des documents d'activité.
- Le chapitre 5 relate des cas pratiques de mise en œuvre.

---

<sup>2</sup> Il existe des expressions francophones telles que : « informatique nuagique » et « infonuagique ». Il sera important de prendre en compte cette terminologie lors des prochains travaux de traduction en cohérence avec les travaux réalisés sur la série des 3030x.

## 2 LE CONTEXTE NORMATIF ACTUEL

Le contexte normatif en cours de développement peut être résumé par le schéma suivant publié dans les normes ISO 30300 et 30301. La réunion qui se tiendra à Paris en juin 2013 apportera des évolutions sur les projets de l'ISO TC46/SC11.



## 3 METADONNEES RAPPELS ET NORMES EXISTANTES

---

L'ISO 30300 précise dans le chapitre 3.1 « Termes relatifs aux documents d'activité », que les métadonnées sont des « données décrivant le contexte, le contenu et la structure des documents d'activité ainsi que leur gestion dans le temps ».

L'ISO 30301 précise dans le chapitre 8.1 « Conceptions des processus liés aux documents d'activité » qu'il faut lors de la création d'un SGDA<sup>3</sup> « déterminer le contenu, le contexte et les informations de contrôle (métadonnées) devant être inclus dans les documents d'activité ».

### 3.1.DEFINITIONS, RAPPELS DES FONCTIONS ET DES SPECIFICITÉS DES MÉTADONNÉES

---

#### 3.1.1. DÉFINITION

---

Les métadonnées (littéralement « données sur/à propos des données ») sont des informations structurées permettant d'identifier et de décrire les documents et/ou données produits dans le cadre d'une activité et de les ré-exploiter à court, moyen et long terme en fonction des besoins des utilisateurs. Ainsi, elles sont indispensables pour restituer la production documentaire d'un point de vue technique mais aussi d'un point de vue intellectuel. Elles permettent donc de rendre l'information accessible et exploitable. En ce sens, les métadonnées sont un outil indispensable à la gestion documentaire.

On pourrait comparer les métadonnées à une étiquette sur une boîte de conserve. Celle-ci est indispensable à l'information du consommateur : elle donne un ensemble de renseignements sur le contenu de la boîte (la marque, le poids, le détail des ingrédients et leur valeur énergétique, la date de mise en boîte, la date de péremption, le nom du producteur/distributeur, son adresse...) permettant ainsi de fournir des informations essentielles. L'« étiquette » d'un document fonctionne de la même manière : elle permet d'identifier aisément son contenu dans son contexte de production en lien avec le processus de travail (ou l'activité) grâce à des informations structurées telles que la date de création, l'entité créatrice, etc. En documentant l'information produite, les métadonnées apportent une valeur ajoutée à l'identification du document. De même qu'on aurait du mal à identifier parmi un lot de boîtes de conserve non étiquetées celle qui contient des petits pois, de même il serait difficile d'identifier parmi des documents dépourvus de métadonnées celui qui correspondrait par exemple au rapport d'activité d'une année donnée.

Dans son article « Histoire(s) de notices »<sup>4</sup>, Gautier Poupeau indique que « l'emploi massif du terme métadonnée constitue un changement sémantique dont il ne faut pas négliger la portée. Ce terme acte le fait qu'on ne s'intéresse plus uniquement à la notice [descriptive accompagnant un document par exemple] dans sa globalité, mais aux différentes portions d'information qu'elle contient à travers la notion de métadonnées ».L'information est ainsi appréhendée à travers ce qui la caractérise au niveau le plus fin : pas de notice descriptive sans champs de notice.

---

<sup>3</sup> SGDA, Système de gestion des documents d'activité – Management system for records

<sup>4</sup> POUPEAU (Gautier), « Histoire(s) de notices » in : *Le document numérique à l'heure du web de données*, séminaire INRIA (Carnac, 1<sup>er</sup>-5 octobre 2012), Paris : ADBS, 2012.

---

### 3.1.2. LES MÉTADONNÉES REMPLISSENT DES FONCTIONS DIFFÉRENTES

---

Tout document numérique doit être accompagné de métadonnées. Plusieurs catégories de métadonnées correspondant à des fonctions différentes et à des niveaux d'exigence variables peuvent être identifiées. Il est possible de distinguer, dans le domaine de la gestion documentaire, les catégories suivantes :

- métadonnées techniques
- métadonnées d'administration
- métadonnées descriptives ou de contenu
- métadonnées de gestion des documents d'activité (records management) qui peuvent comprendre à la fois des métadonnées techniques, descriptives et d'administration.

---

#### LES MÉTADONNÉES TECHNIQUES

---

La première couche de métadonnées se compose tout d'abord d'informations techniques indispensables à une machine pour réinterpréter les fichiers. Le propre de ces métadonnées est d'être générées automatiquement par les systèmes d'information ayant servi à produire les documents/données. On pourrait comparer ces informations minimales aux données de base qui sont renseignées sur une enveloppe (nom, prénom, nom et numéro de rue, ville, code postal) sans lesquelles il serait difficile au facteur de délivrer le courrier à son destinataire.

Cependant, ces métadonnées techniques sont insuffisantes du point de vue de la gestion documentaire. On peut penser par exemple qu'un système de gestion des documents d'activité électroniques ayant pour objectif de garantir l'authenticité et l'intégrité des documents devra intégrer d'autres métadonnées spécifiques telles que : l'horodatage, le calcul de l'empreinte des documents, le type MIME des fichiers vérifié par des outils spécifiques comme Jhove et DROID (Digital Record Object Identification). Il est donc possible de définir un jeu de métadonnées techniques plus complet en fonction des objectifs de gestion qui auront été définis.

---

#### LES IDENTIFIANTS UNIQUES

---

Souvent assimilés à des métadonnées techniques, les identifiants uniques peuvent être considérés comme une typologie à part entière de métadonnées compte tenu de la spécificité de leur rôle. L'identifiant sert de critère diacritique, permettant d'identifier une ressource avec certitude par rapport à un processus d'activité ou bien au sein d'un système d'information.

Les identifiants uniques reposent sur des règles de normalisation qui ne doivent pas être confondues avec de simples règles de nommage de fichiers.

---

#### IDENTIFIER UNE RESSOURCE PAR RAPPORT À UN PROCESSUS D'ACTIVITÉ

---

### **L'exemple du numéro NOR**

Le NOR est un identifiant unique délivré par l'application S.O.L.O.N. pour identifier les projets de textes législatifs et réglementaires ayant vocation à être publiés au Journal officiel de la République française. Il se compose de douze caractères alphanumériques formant cinq ensembles cohérents :

Ex. MAEH1252486D

- Un premier ensemble constitué de trois lettres correspondant au ministère porteur du texte (en référence à une table remise à jour à chaque changement de gouvernement).
- Un deuxième ensemble constitué d'une lettre correspondant à la direction du ministère porteuse du texte (en référence à une table remise à jour à chaque changement de gouvernement).
- Un troisième ensemble constitué de deux chiffres correspondant à l'année de création du texte (ce chiffre est délivré automatiquement par le système et remis à jour chaque année).
- Un quatrième ensemble constitué de cinq chiffres correspondant à une série numérique incrémentée par le système générant le texte lors de chaque création de texte (la série numérique est réinitialisée chaque année).
- Un cinquième ensemble constitué d'une lettre correspondant au type de texte créé (en référence à une table).

### **L'exemple du numéro ECLI5**

Le numéro ECLI (European Case Law Identifier) est un identifiant conçu pour contribuer à la citation correcte et sans équivoque des décisions de justice ayant trait au droit de l'Union Européenne rendues par les juridictions européennes et nationales.

L'identifiant ECLI est composé de cinq éléments :

1. ECLI : Abréviation indiquant la nature de l'identifiant
2. Un code pays ou un code d'organisation : pour la France c'est « FR »
3. L'abréviation de la juridiction
4. L'année de la décision
5. Un numéro d'ordre

Chaque élément est séparé par le caractère deux-points « : »

### LES MÉTADONNÉES D'ADMINISTRATION

On trouve également une catégorie de métadonnées qui permettent d'administrer un système: gérer les profils des utilisateurs et leurs droits.

### LES MÉTADONNÉES DESCRIPTIVES OU MÉTADONNÉES DE CONTENU

Le rôle de ces métadonnées est de décrire le contenu des documents afin d'en faciliter l'accès. Il s'agit sans aucun doute de la catégorie de métadonnées la plus répandue.

---

5 <http://www.parcesmotifs.net/spip.php?article10>

L'effort de normalisation est important en la matière à travers l'exemple des thésaurus et autres formes de langages contrôlés. Dans le domaine de l'indexation descriptive, il n'y a pas de modèle unique : par définition chaque langage d'indexation est créé pour répondre à un domaine d'activité précis.

Quelques applications concrètes :

**EUROVOC**<sup>6</sup> : thésaurus multilingue utilisé par le Parlement européen, l'Office des publications des Communautés européennes, les parlements nationaux et régionaux en Europe, des administrations nationales et certaines organisations européennes. Il couvre tous les domaines d'activité de l'Union européenne et permet d'indexer les documents dans les systèmes documentaires des institutions européennes et de leurs utilisateurs.

**RAMEAU** : pour l'indexation des contenus bibliographiques dans les bibliothèques.

**Thésaurus W**: pour décrire les fonds dans les services d'archives des collectivités territoriales en France.

L'ensemble des catégories présentées ci-dessus sont fondamentales mais insuffisantes pour couvrir le périmètre de la gestion des documents d'activité (records management).

---

### 3.1.3. LA SPÉCIFICITÉ DES MÉTADONNÉES POUR LA GESTION DES DOCUMENTS D'ACTIVITÉ (RECORDS MANAGEMENT)

---

Les métadonnées pour la gestion des documents d'activité sont nécessaires pour conserver et restituer un document dans son contexte de production et préserver ainsi son lien avec le processus de travail (ou l'activité). Elles permettent également de gérer le cycle de vie de l'information (les durées de conservation ou calendrier de conservation, le sort final). Elles contribuent à garantir le caractère authentique des documents (gestion de la preuve).

La norme ISO 23081 décrit un cadre conceptuel général permettant d'identifier les principales fonctions des métadonnées pour la gestion des documents d'activité :

- protection du caractère probant des documents (ce point est particulièrement commenté dans le § 3.2.2)
- garantir leur accessibilité/ exploitabilité dans le temps
- garantir l'authenticité, la fiabilité et l'intégrité des documents dans le temps
- gérer l'accès, les droits et la protection (des données personnelles)
- interopérabilité
- maintenir les liens entre les documents et leur contexte de création par le biais d'une structure fiable
- identifier l'environnement technique de création des documents
- faciliter la migration des documents d'un environnement technique à un autre

---

#### LE DÉFI DE L'INTEROPÉRABILITÉ : L'IMPORTANCE DE LA NORMALISATION

---

Les métadonnées se présentent également comme une réponse aux besoins croissants d'échanges d'informations. En ce sens, l'interopérabilité est devenue un enjeu majeur pour améliorer la gestion de l'information et sa diffusion.

---

<sup>6</sup>[http://europa.eu/eurovoc/sg/sga\\_doc/eurovoc\\_dif!SERVEUR/menu!prod!MENU?langue=FR](http://europa.eu/eurovoc/sg/sga_doc/eurovoc_dif!SERVEUR/menu!prod!MENU?langue=FR)

La normalisation à des fins d'interopérabilité a été très tôt pratiquée par les bibliothèques avec la codification du signalement de leurs documents. Cela s'est traduit concrètement par une structuration des documents sous forme de notices papier comprenant différents champs puis la définition en 1954 d'une norme, l'ISBD (description bibliographique internationale normalisée) visant à établir une cohérence lors du partage de l'information bibliographique. Pour ce faire l'ISBD « détermine quels éléments de données doivent être notés ou transcrits en une séquence spécifique qui sert de base à la description de la ressource en cours de catalogage »<sup>7</sup>. Avec la généralisation des catalogues en ligne, de nouvelles normes concernant les champs de notices ont ensuite vu le jour (formats MARC).

La normalisation de la description des fonds dans les services d'archives a été beaucoup plus tardive : ce n'est qu'en 1994 que la norme ISAD(G) (norme générale et internationale de description archivistique) a été publiée et il a fallu attendre 2002 pour qu'on voie apparaître une version stabilisée de la DTD EAD (description archivistique encodée structurée avec le langage XML). L'EAD est en cours de révision actuellement pour faire évoluer la DTD vers un schéma XML qui facilitera l'interopérabilité avec d'autres formats de métadonnées.

Plusieurs tentatives de modélisation des métadonnées ont eu lieu depuis les années 1990<sup>8</sup>:

- 1996 : le CIDOC-CRM (*Committee on documentation – Conceptual reference model*) : fournir un cadre conceptuel de référence pour décrire toutes les interactions d'un objet avec d'autres entités dans un modèle orienté objet ;
- les FRBR (*Functional Requirements for Bibliographic Records*) publiés en 1998 : définissent trois groupes d'entités pour décrire les documents, les personnes physiques et morales et les différents concepts qui peuvent être sujets des œuvres ;
- le *Dublin Core* qui définit en 2007 un modèle abstrait (DCAM, *Dublin Core Abstract Model*).

Les métadonnées pour la gestion des documents d'activité ont connu un développement encore plus tardif. En effet, la norme ISO 23081 qui présente un cadre conceptuel pour la gestion de ces métadonnées, a été publiée entre 2004 et 2006. Des normes comme MoReq et l'ISO 16 175 (ICA-Req) ayant pour objectif de fournir des lignes directrices avec des modèles pratiques pour la gestion des documents d'activité ont complété ensuite le champ normatif sans qu'on dispose pour autant d'un modèle unifié et standardisé pour les métadonnées.

Toutefois, la norme ne préconise pas de modèle de métadonnées précis, elle ne fournit pas une check-list. Actuellement, il n'y a pas de modèle francophone en termes d'exigences minimales et de vocabulaire homogène.

Il faut un modèle qui permette de faire converger les différentes initiatives existantes éparse ainsi que les normes existantes, à la fois l'ISO 15489, l'ISO 30300, l'ISO 23081 et l'ISO 16175 (ICA-Req), sans oublier qu'il ne peut y avoir qu'un socle commun minimal car chaque processus d'activité est par définition unique.

---

<sup>7</sup> [http://www.bnf.fr/documents/isbd\\_integre.pdf](http://www.bnf.fr/documents/isbd_integre.pdf)

<sup>8</sup> Cf. Gautier Poupeau

Cet enjeu devient prépondérant dans l'administration publique : avec l'effort de rationalisation des finances publiques mené par les réformes récentes (LOLF<sup>9</sup>, RGPP<sup>10</sup>, modernisation de l'action publique) la nécessité de davantage de transparence, de traçabilité et d'interopérabilité (ex. interopérabilité des systèmes pour faciliter l'archivage électronique définitif des métadonnées) s'est affirmée.

### 3.2. LES NORMES EXISTANTES, LEUR LIMITES ET LEUR PERTINENCE

Les jeux de métadonnées sont pléthoriques. Il en existe pratiquement pour chaque cas particulier de données ; pour qualifier les formats, pour quantifier les volumes, pour évaluer les échanges, les flux, pour identifier les "originaux" des "copies", pour tracer la circulation des données dans l'entité productrice, etc. Or, singulièrement, les jeux de métadonnées propres au *records management* n'ont pas encore été stabilisés dans leur forme. Des recommandations existent, mais elles sont souvent inadaptées, car trop génériques.

#### 3.2.1. POSITIONNEMENT DES NORMES ISO 23081, ISO 15489 ET 30300

Les métadonnées jouent un rôle prépondérant dans le processus de suivi et de contrôle du cycle de vie des documents d'activité et dans la traçabilité des actions dans, autour et sur les données et *records* : les normes de *records management* accordent une attention particulière à la description des contenus et des structures des jeux de métadonnées. Ces métadonnées ont donc leur importance dans des projets de *records management*, pour garantir l'identification, la fiabilité, l'intégrité et l'authenticité des documents d'activité, ainsi que l'ensemble des moyens matériels, humains et financiers des activités et processus qui président à leur création et à leur existence (« *records management metadata can be used to identify, authenticate and contextualize records and the people, processes and systems that create, arrange, maintain and use them and the policies that govern them [...] metadata ensure authenticity, reliability, usability and integrity over time and enable the management and understanding of information objects* », norme ISO 23081-1 (partie 4)).

La norme ISO 15489-1 définit les métadonnées propres au *records management* comme des données décrivant le contexte, le contenu et la structure des documents d'activité ainsi que leur gestion dans la durée (ISO 15489-1 (paragraphe 3.2) reprise dans les normes ISO 23081-1 (paragraphe 4), ISO 30300 et 30301 et approfondie dans la norme ISO 16171 (4312)). Cette définition est ensuite reprise dans la norme ISO 23081-1 (partie 5.2.1) qui distingue deux sortes de métadonnées :

- les métadonnées de contexte, de contenu, d'apparence et de structure utilisées pour la création ou la capture du document d'activité,
- les métadonnées de gestion des documents d'activité et des processus par lesquelles le document est utilisé, y compris les changements induits sur le contenu, la structure et l'apparence.

Le mode de gestion de ces métadonnées doit être pérenne dans le temps et ce quel que soit l'environnement (ISO 15489-1, 16171 et 23081-1).

---

<sup>9</sup>LOLF : Loi organique relative aux lois de finances.

<sup>10</sup>RGPP : Révision générale des politiques publiques.

Toutes les normes citées insistent plus ou moins de manière détaillée sur la définition du jeu de métadonnées dans le processus de mise en place d'une politique de *records management* (ISO 15489 partie 7.1 et ISO 23081-1 (partie 5.2.1), sur la structuration du set (ISO 15489-1 partie 9), et sur la gestion du set (ISO 23081-1, partie 5.2.3 et ISO 16171 termes et définitions, et ISO 15489-2 (partie 4.3.1.)).

Cependant, si ces quatre normes se complètent relativement bien pour tenter de définir des exigences en matière de métadonnées, elles ne recommandent pas de jeu-type adaptés aux documents d'activité, ou de grille d'évaluation de jeux de métadonnées existants dans des systèmes de gestion de documents d'activité.

### 3.2.2. L'ISO 23081-1 ET LA NOTION DE VALEUR DE PREUVE

Début 2012 a été soumise à enquête publique par l'AFNOR la traduction de l'ISO 23081-1 - Information et Documentation – Processus de records management – Métadonnées pour les records – partie 1 : Principes. Cette norme ne pouvait être révisée sur le fond puisque déjà publiée à l'ISO en anglais en 2006, l'objet de l'enquête était donc de savoir si l'on souhaitait disposer de la traduction française telle que soumise à enquête.

La principale objection soulevée a porté sur ce qui a été qualifié d'erreur de traduction, qui en y regardant de plus près tenait à l'ambiguïté de la rédaction anglaise. La traduction incriminée se trouvait dans la section « 5.1 Finalités et bénéfices des métadonnées de la gestion des documents d'activité<sup>11</sup> ». Citons les termes de l'objection : le texte « laisse penser qu'il suffit que les « documents d'activité (*records*) » soient accompagnés de métadonnées pour qu'ils aient « valeur de preuve ». Ce qui constitue une confusion très gênante entre la notion de 'condition nécessaire' et de 'condition suffisante'. » Soit pour l'exprimer autrement : toutes les métadonnées ne confèrent pas une valeur de preuve, seules certaines le font, et dans tous les cas à elles seules les métadonnées ne suffisent pas à garantir cette valeur.

Concrètement, la traduction initialement proposée et l'ambiguïté du texte anglais rendaient et rendent toujours cette norme difficilement voire inutilisable en l'état dans le cadre d'activités commerciales autour de l'archivage électronique, car incompatible avec la norme de référence dans ce domaine, la NF Z42-013: 2009 (et son équivalence internationale ISO 14641-1: 2012).

D'un point de vue légal, on peut aussi interroger la pertinence et la solidité de la norme ISO 23081-1 telle qu'actuellement publiée. Car si « certaines » métadonnées pourraient conférer une valeur de preuve, la norme n'aide pas à construire concrètement un jeu de métadonnées de référence – les recommandations restent générales et n'apportent pas grand-chose.

Un consensus français avait été obtenu lors de la réunion de dépouillement qui s'était tenue en juillet 2012, en précisant que « certaines métadonnées contribuent à assurer » la valeur probante. Mais au-delà de cette « béquille », les commentateurs présents ce jour-là incitaient fortement, pour les raisons évoquées, à une révision de la norme initiale.

<sup>11</sup> « metadata support business and records management processes by protecting records as evidences » traduit par « les métadonnées (...) confèrent aux documents leur valeur de preuve. »

### 3.2.3. LA NORME ISO 15836 OU DUBLIN CORE

Un jeu de métadonnées que l'on pourrait croire adapté au *records management* serait le *Dublin Core* : objet de la norme ISO 15836, de la norme ANSI/NISO Z39.50, reposant sur les standards du W3C, largement utilisé par les systèmes de gestion d'informations dans tous les domaines et pas seulement ceux des bibliothèques. Le *Dublin Core* est un jeu de métadonnées destiné fondamentalement à la recherche documentaire, et repose sur quinze métadonnées, ou éléments, qu'on peut qualifier (c'est à dire restreindre la définition des éléments) ou non. Notons au passage que la norme ISO 1583 ne s'attache qu'à décrire les métadonnées et leur donne une traduction française, sans méthode de mise en œuvre. Or, le but premier du *Dublin Core* est bien la recherche documentaire, et ce but n'est qu'un des aspects d'une politique de *records management*.

Paradoxalement, la très large utilisation de ce jeu de métadonnées, par une variété impressionnante d'organisations, doit nous alerter : n'y a-t-il pas une marge importante de libre appréciation laissée à l'organisation pour remplir le contenu de chaque métadonnée ? Par exemples, le contenu de la métadonnée "date" peut être aussi bien être rempli par un organisme selon la date de création du document, alors qu'un autre établissement choisira de n'indiquer que la date de dernier enregistrement. L'absence de concept clair d'horodatage ou d'identifiant-unique est une source de risque dans la mise en œuvre des systèmes.

Ainsi, pour un jeu de métadonnées de *records management*, et en suivant les lignes directrices développées dans les normes 15489 et 30300, les quinze items constituant le *Dublin Core* ne sont pas antagonistes, mais sont insuffisants. Le *Dublin Core* ne peut être appliqué « tel quel » à la gestion des documents d'activité et devrait être complété et adapté.

### 3.2.4. FOCUS SUR L'IDENTIFIANT UNIQUE

Un des éléments particulièrement représentatifs de la réflexion préalable à la constitution d'un jeu de métadonnées porte sur l'identifiant du document d'activité. Si dans certaines entreprises ou organisations, une nomenclature pour le nommage de documents a été mise en place dans le cadre d'une démarche de *records management*, il en va tout autrement dans le contexte électronique avec les métadonnées.

La partie 9 de la norme ISO 15489-1 indique que l'identifiant est un élément incontournable des métadonnées de description. Il doit être unique, et permet de référencer le document dans le plan de classement de l'organisation. La norme ISO 15489-2 précise que l'identifiant est généré dès l'intégration du document d'activité dans le système de gestion.

Pour parvenir à une mise en conformité de ces exigences normatives, plusieurs solutions existent, entre les solutions de nommage "maison" ou le recours à des systèmes d'identifiants par exemple l'ISSN ou le DOI (*Digital Object Identifier*) élaborés par des tiers, appelés autorités « nommantes » ("*namingauthority*").

## 4 REVERSIBILITÉ, SÉCURITÉ, CLOUD : QUELLES EXIGENCES IMPACTANT LES MÉTADONNÉES ?

---

### 4.1. EXIGENCE DE RÉVERSIBILITÉ

---

La réversibilité est un processus permettant de récupérer des données contenues dans un système d'information afin de les transférer vers un autre système. Cette opération est également appelée migration. Elle est nécessaire lors du remplacement d'un système ou d'un changement de prestataire assurant la conservation des données.

La sauvegarde se distingue de la migration car elle est avant tout prévue pour restaurer un système à l'identique après un incident ayant endommagé les données du système.

Dans le domaine documentaire, la migration concerne des objets numériques (documents, dossiers, photos, ...) et les métadonnées qui caractérisent ces objets.

Trois éléments sont nécessaires pour qu'un système soit réversible :

- La source doit disposer d'une fonction permettant de sélectionner et d'exporter les données, par exemple des documents accompagnés de leurs métadonnées ;
- la cible doit disposer d'une fonction permettant d'importer les données et de les exploiter ;
- un format de transfert compatible avec la source et la cible doit être spécifié.

En dehors du cas où le système source et le système cible proviennent d'un même éditeur assurant la migration des données, il est nécessaire de choisir et de paramétrer des formats normalisés tels que ISO 22938 (Gestion de documents - Format d'échange de données pour la gestion de documents/contenu électroniques) ou ISO 2709 (Format d'échange pour les données bibliographiques).

Quel que soit le type d'application, l'existence d'une réversibilité opérationnelle est un élément clé pour la pérennité et la sécurité des données conservées au fil du temps dans un système.

### 4.2. EXIGENCE DE SÉCURITÉ

---

La sécurité des SGDA est bien évidemment liée à celle des métadonnées associées aux documents d'activité. C'est la raison pour laquelle ce Livre blanc a souhaité porter une attention particulière au couple métadonnées et sécurité eu égard aux normes les plus récentes en matière de sécurité.

Les métadonnées pourraient permettre de disposer d'éléments relatifs à la sensibilité et aux besoins de sécurité associés à une information. Cela permettrait de faciliter le traitement de cette information, relativement à la sécurité, tout au long de son cycle de vie.

Les éléments nécessaires découlent de l'analyse de risque associée à cette information. Ils peuvent concerner les critères classiques de l'analyse de risque en sécurité de l'information (disponibilité, intégrité, confidentialité), mais aussi ceux qui auront été traités de façon complémentaire lors de cette analyse (imputabilité, authenticité, pérennité, etc.). Ils peuvent aussi, plus globalement, se référer au niveau de sensibilité de l'information tel qu'il sera défini en fin d'analyse.

Le JTC1 est, à l'ISO, le Comité Technique qui gère les technologies de l'information. Au sein du JTC1, le SC27 est un sous-comité dédié à la sécurité. Le SC27 se compose de 5 *Working Groups* parmi lesquels le WG1 et le WG4 sont en charge de l'élaboration des normes de la

série 27000 construites autour de la notion d'ISMS (*Information Security Management System*).

Les trois normes centrales de cette série sont :

- ISO/IEC 27001, norme d'exigence qui permet l'analyse de conformité et la certification en sécurité de l'information.
- ISO/IEC 27002, norme de bonnes pratiques associée à la précédente et qui rassemble 130 mesures.
- ISO/IEC 27005, norme sur le management du risque en sécurité de l'information qui vient en soutien méthodologique des deux précédentes.

La série 27000 se compose actuellement d'un peu plus de 30 normes.

Les développements les plus récents sont les suivants :

Ces normes ci-dessous représentent d'une part les évolutions récentes au sein de la série 27000 et d'autre part les développements en cours concernant le *Cloud Computing*. La série 27000 comporte actuellement une trentaine de normes dans le contexte de la norme certifiante ISO 27001.

- ISO/IEC 27035 - Information security incident management – Août 2011. C'est une norme "bottom-up" qui traite de la gestion d'incident en sécurité de l'information.
- ISO/IEC 27037 - Guidelines for identification, collection, acquisition, and preservation of digital evidence – Octobre 2012

Cette norme vise à permettre de garantir l'intégrité des preuves numériques pour en autoriser l'exploitation dans des cadres légaux. Elle introduit des critères complémentaires concernant les précautions à prendre dans la gestion des données, notamment : auditabilité, répétabilité, reproductibilité et justifiabilité.

#### NORMES ISO/IEC 270XX LIÉES AU CLOUD

Trois normes concernant le *Cloud Computing* sont actuellement en préparation au SC 27 :

- ISO/IEC 27017 - *Guidelines on information security controls for the use of cloud computing services based on ISO/IEC 27002* – WD 3

Cette norme vise à permettre l'extension de l'utilisation des normes 27001/27002 au cas du *Cloud Computing*.

- ISO/IEC 27018 - *Code of practice for data protection controls for public cloud computing services* – WD2.

Cette norme est élaborée au WG 5 (*Privacy technologies and Identity Management*). Elle concerne la protection des données et plus particulièrement des informations personnelles.

- ISO/CEI 27036 – *Part 5 : Information security for supplier relationships* - WD

Ce projet est actuellement en cinq parties (*Overview and Concepts, Common Requirements, Guidelines for ICT Supply Chain security, Outsourcing, Guidelines for security of cloud services*). La cinquième partie, qui concerne la sécurité des services dans le *Cloud Computing*, est en WD.

### 4.3.COMMENTAIRES VIS-À-VIS DU CLOUD STORAGE

---

L'expression *Cloud Computing* ou Informatique en nuage est apparue il y a quelques années comme concept dans la littérature scientifique et comme offre commerciale proposé par les fournisseurs de services informatiques. Vu comme un nouvel Eldorado cette expression a souvent été utilisée sans trop de précaution.

Pour sa partie NIST, *National Standard Institut* retient la définition suivante : « Le *Cloud Computing* est un modèle informatique qui permet un accès facile et à la demande par le réseau à un ensemble partagé de ressources informatiques configurables (serveurs, stockage, applications et services) qui peuvent être rapidement provisionnées et libérées par un minimum d'efforts de gestion ou d'interaction avec le fournisseur du service ».

Dans ce contexte de ressources partagées et très souvent externalisées, le *Cloud Storage*, prend une dimension particulière dès lors qu'il s'agit de stocker, voire de conserver des documents d'activité. Si le *Cloud Storage* a de nombreux avantages, en particulier économique il a aussi un certain nombre d'aspects vis-à-vis desquels il faut avoir une vigilance avisée comme par exemple l'intégrité, la sécurité, la réversibilité, la portabilité, la résilience ou continuité de service.

En conséquence, deux remarques peuvent être faites. La première est que le *Cloud Storage* est une solution d'ordre technique. Elle ne relève pas de considérations fonctionnelles et en particulier de celles concernant les métadonnées. En revanche, et c'est la seconde remarque, il est impératif pour toutes organisations qui envisagent des prestations de *Cloud Storage* soient attentives aux aspects contractuels qui la lie à un fournisseur en matière de sécurité, de pérennité, de protection des données et de réversibilité.

## 5 ÉTUDES DE CAS ET POINTS DE VUE

---

### 5.1.LINKLATERS, CABINET D'AVOCAT INTERNATIONAL(SYLVIE FORASTIER - LUXEMBOURG)

---

#### 5.1.1. OBJECTIF : RÉPONDRE AUX QUESTIONS PRINCIPALES D'IDENTIFICATION ET DE LOCALISATION DES DOSSIERS.

---

Pour ce faire, il faut distinguer les données, des métadonnées. Les données sont ce qui va être d'utilité immédiate à l'utilisateur pour identifier le dossier et le localiser (exemple : nom du dossier, forme physique, le nombre d'occurrence, localisation, statut etc.). Selon la nature des métadonnées, elles ne vont pas être sollicitées lors de chaque recherche mais elles vont servir à affiner ou à préciser la recherche quand celle-ci devient complexe. Elles vont également servir à assurer la traçabilité de tout le cycle de vie d'un dossier donné. Dans un univers où le dossier est protéiforme et est constitué de documents papier, de documents électroniques, d'emails et de base de données, les métadonnées vont être de nature suivante:

#### 5.1.2. MÉTADONNÉES SUR LES DOCUMENTS PAPIER/ÉLECTRONIQUES:

---

Elles sont pour la plupart créées lors de la production d'un document. Certaines sont générées automatiquement comme le numéro du document, la date de création, etc. D'autres sont manuelles (comme la référence du dossier); Elles sont destinées principalement à l'usage du public et de l'administrateur bien que chacun va en faire un usage différent; L'administrateur va identifier la source du document grâce à ses références (par exemple : la date, les initiales du créateur, le numéro du dossier, etc.) et

définir la destination du document pour alimenter le dossier adéquat. L'utilisateur (qui peut être différent du créateur) va quant à lui se servir des mêmes métadonnées dans un objectif de recherche ; elles lui serviront à vérifier qu'il a atteint son but et trouvé le document requis.

Ces métadonnées sont codifiées de façon uniforme au sein de l'entreprise (emplacement sur le document, abréviations, etc.) ce qui en rend la gestion assez facile. Toutefois l'exploitation de ces métadonnées repose sur leur exactitude, en particulier pour ce qui concerne les métadonnées à encoder manuellement.

---

### 5.1.3. MÉTADONNÉES SUR LES DOSSIERS PHYSIQUES:

---

Elles sont générées par la base de données (voir métadonnées standards - dans la base de données) mais l'administrateur opère une sélection parmi les plus significatives car toutes ne peuvent figurer sur le dossier physique (risque de bruit : seules des métadonnées comme le code barre, le numéro d'archives vont être pertinentes à reporter sur le dossier physique). A la place des métadonnées écartées, l'administrateur va en ajouter d'autres porteuses de sens (par exemple un code couleur). Comme pour les métadonnées des documents, l'usage qui en sera fait va varier entre l'administrateur et l'utilisateur à savoir que l'administrateur va être dans une démarche de mise à disposition des dossiers tandis que l'utilisateur sera dans la position de rechercher les dossiers.

Ces métadonnées sont codifiées par l'administrateur : il a donc fallu déterminer les métadonnées les plus utiles à l'utilisateur. Cette étape peut prendre du temps puisqu'elle repose sur une connaissance des processus métier. Le contrôle opéré sur ces métadonnées doit être rigoureux et clairement communiqué car elles sont susceptibles d'évolution (par exemple un ajout dans le code couleur) et elles véhiculent une information: un mauvais usage de ces métadonnées risque de créer une mauvaise information.

---

### 5.1.4. MÉTADONNÉES DE LA BASE DE DONNÉES:

---

#### MÉTADONNÉES STANDARDS

---

Ces métadonnées sont des métadonnées standards car ce sont celles que l'on va retrouver quel que soit l'outil.

Elles sont créées par l'outil lui-même et ni l'administrateur, ni le public ne peut influencer sur celles-ci (code barre, numéro d'archive, date de création, date d'archivage, information migration, historique des emprunts, etc.). D'ailleurs l'usage que l'utilisateur peut faire de ces métadonnées est très restreint et comme l'accès à certaines métadonnées lui est impossible, il sera obligé d'adresser une requête à l'administrateur en cas de besoin. Ces métadonnées sont en majorité à l'usage de l'administrateur qui va d'une part décider de celles disponibles pour les utilisateurs (voir métadonnées des dossiers physiques) et d'autre part se servir de ces informations dans un but d'administration générale du service de Records Management.

Ces métadonnées sont les plus faciles à gérer car elles sont codifiées et créées par l'outil lui-même sans aucun risque d'erreur.

#### MÉTADONNÉES PERSONNALISÉES

---

Outre les catégories de métadonnées citées ci-dessus, il existe dans la base de données, des métadonnées personnalisées qui sont liées aux besoins spécifiques de l'entreprise. Créées par l'administrateur, elles ne sont pas générées de façon standardisée mais uniquement au cas par cas et doivent être accessibles par les utilisateurs quand ceux-ci estiment en avoir besoin (exemple: information comptable).

Il existe également des métadonnées personnalisées à l'usage administratif du service de records management (exemple : agent opérateur de l'archivage).

Ces métadonnées sont codifiées par l'administrateur : sur base de la connaissance des processus métier, il a préalablement fallu déterminer les métadonnées les plus utiles aux utilisateurs, décider d'une codification et les encoder de manière à ce que l'utilisateur puisse y accéder.

#### MÉTADONNÉES DE COMMUNICATION

---

Elles sont créées par les utilisateurs (par exemple: email, bordereau de versement etc.) et sont à destination de l'administrateur. L'objectif de ces métadonnées est purement administratif et va servir à la gestion générale du service de Records Management (demande de création, ordre d'archivage, etc.). L'objectif principal est d'assurer la traçabilité des instructions données.

Elles sont très difficiles à gérer car elles sont de sources multiples et existent en grand nombre. Elles sont par conséquent difficile à exploiter (comment retrouver un mail au milieu de 5000 ordres d'archivage ?). Une solution proposée est de compiler ces informations afin qu'elles soient disponibles dans une seule et même source. Par exemple, compiler les ordres d'archivage dans une fiche Excel permet de retrouver grâce à la fonction « filtre » les instructions données et permet ainsi de remonter plus facilement et rapidement à l'instruction d'origine.

#### MÉTADONNÉES DE MIGRATION

---

En cas de migration d'une base de données vers une autre, les métadonnées donnent à l'administrateur la possibilité de vérifier la source et la nature des informations avant la migration. Il s'agit de conserver une copie des données d'origine (avant migration) afin de garantir que la migration a été effectuée correctement. L'objectif de la conservation de ces métadonnées est à l'usage de l'administrateur pour la gestion générale du service de *records Management*.

Les formes et sources variées des métadonnées imposent des modes de gestion différenciés. Il n'existe pas forcément de lien entre toutes les métadonnées. Néanmoins toutes ces métadonnées coexistent pour un dossier donné. De plus, Les métadonnées doivent être enregistrées dans des champs que, soit l'administrateur, soit l'utilisateur, ou les deux, peuvent interroger. Il ne faut pas sous-estimer l'importance des métadonnées car elles sont l'assise de tout système de gestion de données performant.

### 5.2.POINT DE VUE SUR LES MÉTADONNÉES EN ENTREPRISE (LUCAS COLET - LUXEMBOURG)

---

Dans le cadre d'une utilisation opérationnelle, le référentiel de métadonnées de l'OAIS vient assez naturellement comme une première solution pour construire son jeu de métadonnées et les classer en catégories. Néanmoins, dans la pratique, nous pouvons nous rendre compte que trois catégories prédominent :

- Les métadonnées techniques (souvent appelées « de gestion »), qui regroupent toutes les métadonnées ayant trait à la pérennisation de l'information, à son support, à son sort final, et à son utilisation et accès, qui seront donc utilisées majoritairement par le responsable (logiciel ou non) du stockage de l'information. Elles peuvent également servir à la création d'un environnement de preuve. Ces métadonnées techniques peuvent s'appuyer sur MoReq2, qui dresse un panorama relativement complet de celles-ci ;
- Les métadonnées métier (souvent appelées « descriptives »), qui sont propres au cœur de métier et qui facilitent l'identification et la recherche dans un contexte

donné (un service RH aura des métadonnées différentes à ce niveau que le service Facturation, par exemple) ;

- Les autres métadonnées, qui permettent de faciliter la présentation de l'information, sa recherche, et tout ce qui n'est pas couvert par les deux autres catégories.

Néanmoins, dans le cadre de la construction du jeu de métadonnées, il reste pertinent de s'appuyer sur une norme pour permettre une interopérabilité maximale de ces métadonnées.

### 5.3. LES MÉTADONNÉES ET LEUR RÔLE DANS LA GESTION ÉLECTRONIQUE DES DOCUMENTS – UN APERÇU DE L'EXPÉRIENCE DE LA COMMISSION EUROPÉENNE (MATTIAS HALLIN)

---

La Commission européenne a entamé, en 2002, la modernisation de la gestion et de l'archivage de ses documents<sup>12</sup>, en vue d'accroître l'efficacité et la transparence de ses travaux et de garantir qu'elle puisse à tout moment rendre compte de ce dont elle est redevable.

Cette opération s'est notamment traduite par l'adoption par la Commission de deux décisions qui annexent à son règlement intérieur des dispositions concernant, d'une part, l'administration des documents<sup>13</sup> et, d'autre part, les documents électroniques et numérisés<sup>14</sup>. Ces deux décisions ont été complétées entre 2003 et 2008 par des modalités d'application (Mda, refondues en 2009 en un seul texte<sup>15</sup>), ainsi que par une liste commune de conservation (LCC, révisée en 2012)<sup>16</sup>.

Ces modalités d'application comprennent une liste d'une cinquantaine des métadonnées<sup>17</sup> (obligatoires ou facultatives), couvrant la gestion de tout le cycle de vie des documents, des dossiers et du plan de classement de la Commission. La liste est sous-divisée en six groupes : enregistrement des documents<sup>18</sup>, rubriques du plan de classement<sup>19</sup>, dossiers<sup>20</sup>, conservation des dossiers<sup>21</sup>, transfert<sup>22</sup> et élimination<sup>23</sup>.

---

<sup>12</sup>Dans ce texte nous n'utilisons ni « document d'activité » ni « record », mais seulement « document », afin de respecter la terminologie du cadre normatif de la Commission (où ce dernier est le seul terme utilisé, aussi bien en français qu'en anglais, bien qu'il corresponde dans la pratique de très près au terme « document d'activité »).

<sup>13</sup>Décision 2002/47/CE, CECA, Euratom du 23 janvier 2002, JO L 21, du 24.1.2002, p. 23.

<sup>14</sup>Décision 2004/563/CE, Euratom, du 7 juillet 2004, JO L 251, du 27.7.2004, p. 9.

<sup>15</sup>« Modalités d'application de la décision 2002/47/CE, CECA, Euratom concernant l'administration des documents et de la décision 2004/563/CE, Euratom concernant les documents électroniques et numérisés », SEC(2009)1643 du 30 novembre 2009.

<sup>16</sup>« Liste commune de conservation des dossiers au niveau de la Commission européenne », SEC(2007)970 du 4 juillet 2007 et « Liste commune de conservation des dossiers au niveau de la Commission européenne – première révision », SEC(2012)713 du 17 décembre 2012.

<sup>17</sup>SEC(2009)1643, point IV.1.

<sup>18</sup>Les métadonnées d'un document sont : la date du document; la date d'enregistrement; le numéro d'enregistrement (identifiant unique); l'auteur du document et, s'il s'agit d'une personne physique, le service/organisme dont il relève; le destinataire du document et, s'il s'agit d'une personne physique, le service/organisme dont il relève; s'il y a lieu, le niveau de classification du document, ainsi que le timbre et l'identifiant de sécurité; le titre du document ou son objet; le type de document concerné; la rubrique concernée du plan de classement et plus spécifiquement le titre du dossier; dans la mesure du possible, le nombre de pièces jointes et une brève description de celles-ci.

Parmi les sources d'inspiration pour cette liste (qui vise à répondre au mieux aux besoins de l'institution) et pour sa mise en œuvre figurent la norme ISO 15489 ainsi que MoReq, MoReq2 et MoReq2010, bien qu'il a été nécessaire d'interpréter ces sources à la lumière de l'expérience spécifique de la gestion documentaire à la Commission et d'en tenir compte.

Cette liste est applicable quel que soit le support, numérique ou non, mais dans la pratique le rôle des métadonnées est considérablement élargi dans le contexte du numérique. Bien entendu, elles servent en premier lieu à décrire et identifier des objets (comme des documents, des dossiers etc.) d'une façon compréhensible par les utilisateurs. Mais, dans un système de gestion électronique de documents (GED), elles sont aussi directement utilisées par le système lui-même, en permettant d'automatiser des opérations et d'appliquer des règles. Le système électronique peut aussi faire interagir des métadonnées ou en créer de nouvelles à partir de celles déjà existantes.

Outre le cadre normatif, la Commission s'est également dotée d'une famille de systèmes informatiques, appelée Hermes-Ares-Nomcom (HAN)<sup>24</sup>, pour sa gestion des documents.

---

<sup>19</sup>Pour chaque rubrique du plan de classement, les métadonnées sont les suivantes : l'intitulé; une brève description de sa portée; le code; la date de création; la date de clôture; l'état; le service propriétaire et, le cas échéant, les services associés.

<sup>20</sup>Chaque dossier doit contenir les métadonnées suivantes : le code de la rubrique du plan de classement; le code du dossier; le titre du dossier; le chef de file du dossier, y compris, le cas échéant, les chefs de file successifs; s'il y a lieu, le niveau de classification du dossier; la date de sa création; la date de sa clôture; l'état du dossier; s'il y a lieu, le contenu d'un dossier-type; s'il y a lieu, l'application de gestion électronique de documents.

<sup>21</sup>Les métadonnées de conservation des dossiers sont les suivantes : la catégorie de dossier; s'il y a lieu, son contenu; la durée d'utilité administrative (DUA); l'action à effectuer à l'échéance de la DUA; l'action post-transfert aux Archives historiques; le chef de file « conservation »; la localisation; le support; s'il y a lieu, la dé-classification; la liste des documents contenus dans le dossier.

<sup>22</sup>Suite au premier tri, en plus des métadonnées acquises, les métadonnées suivantes doivent être ajoutées à toutes les archives transférées: la date du transfert; l'identification du service qui procède au transfert; le niveau de classification; les exceptions visées à l'article 2 du règlement n°354/83 du Conseil, tel qu'il a été modifié par le règlement n° 1700/2003 du Conseil; les dates d'expiration de la classification ou des exceptions visées à l'article 2 du règlement n° 354/83 du Conseil, tel qu'il a été modifié par le règlement n° 1700/2003 du Conseil (s'il y a lieu).

<sup>23</sup>Suite au premier ou au deuxième tri, les métadonnées acquises suivantes doivent être conservées à titre permanent dans les références des dossiers éliminés: tous les codes de dossier (officiel, spécifique à la DG ou autre); le titre du dossier; le(s) service(s) propriétaire(s)/chef(s) de file; la date de la création; la date de la clôture; la catégorie; la description du dossier, y compris le nombre et le type d'unités à archiver pour les dossiers papier ou la mention «dossier électronique» pour les dossiers électroniques ou la mention «mixte» pour les dossiers composés à la fois de documents papier et de documents électroniques; auxquelles il convient d'ajouter les métadonnées suivantes: la date d'élimination; le service responsable de l'élimination.

<sup>24</sup>Un aperçu de HAN se trouve dans le livre blanc «ISO 30300-30301 – Système de gestion des documents d'activité : définition, modélisations, intégration aux autres normes de système de management», AFNOR, mai 2012, pp. 34 f., disponible au <http://www.bivi.fonctions-documentaires.afnor.org/content/download/32987/197785/version/10/file/AFNOR->

Ceci a demandé une analyse plus approfondie des métadonnées prescrites par le cadre normatif et de leur utilisation afin d'en tirer le plus grand profit à la lumière, précisément, de leur finalité et de leur fonctionnement élargi dans un système de GED.

Il n'est pas possible d'énumérer ici tous les aspects de ce travail dans le détail, mais à titre d'exemple considérons les métadonnées et fonctions liées à la notion de « date de clôture » d'un dossier.

Selon les règles de la Commission la date de clôture est une métadonnée obligatoire de chaque dossier, définie par les MdA comme étant « la date à laquelle le document le plus récent a été ajouté au dossier »<sup>25</sup>. Or, cette date de clôture n'est pas nécessairement la date à laquelle l'opération de clôture du dossier dans le système a été effectuée. Cette dernière date constitue en fait une métadonnée supplémentaire qui n'est pas prescrite par les règles, mais est automatiquement ajoutée par le système dès qu'un utilisateur clôture un dossier (elle est appelée « date de fermeture » afin d'éviter le mieux possible une confusion des deux métadonnées).

Pour connaître la date de clôture, le système doit alors prendre en considération les métadonnées de tous les documents qui sont classés dans le dossier. Ces documents peuvent avoir deux statuts : ou bien ils ne sont que sauvegardés, ou bien ils sont enregistrés.

Dans le contexte actuel un document sauvegardé est un document qui ne remplit pas les critères pour être formellement et obligatoirement enregistré, mais qu'on a quand même décidé de garder dans le dossier afin de le compléter. Pour un document en état de sauvegarde, seule la métadonnée du titre/objet doit obligatoirement être introduite par l'utilisateur pendant que le système ajoute aussi un numéro de sauvegarde et la date de création/sauvegarde. Un document enregistré doit en revanche être accompagné par toutes les métadonnées obligatoires de la liste susmentionnée, en incluant plusieurs dates : la date de création/sauvegarde, la date du document et la date d'enregistrement.

Au moment de déterminer la date de clôture d'un dossier, le système doit alors comparer toutes ces dates et choisir la plus récente, qui est donc copiée dans le champ « date de clôture ». À partir de cette date, et en interaction avec d'autres jeux de métadonnées, le système procède par la suite au calcul de toute une série de métadonnées de conservation pour le dossier.

La LCC susmentionnée établit en fait une durée d'utilité administrative (DUA) ainsi que les actions post-DUA à prendre pour chaque catégorie de dossiers. Ces données ont été encodées dans le système sous la forme d'une liste de catégories LCC, ce qui permet à la fois de consulter la LCC directement dans le système de GED et d'utiliser de façon (semi) automatique ces informations. À chaque rubrique du plan de classement sont liées les catégories pertinentes de la LCC, parmi lesquelles, pour chaque dossier de la rubrique, une

seule est choisie (s'il n'y a qu'une catégorie disponible, elle est automatiquement sélectionnée).

Une fois la date de clôture du dossier calculée comme décrit ci-dessus, le système peut ainsi procéder directement au calcul et aux ajouts, entre autres des métadonnées de dossier suivantes : la date de fin de la DUA<sup>26</sup>, l'action post-DUA à entreprendre (transfert aux archives historiques, éventuellement après échantillonnage ou sélection, ou élimination) ainsi que l'action post-transfert (préservation permanente ou échantillonnage/sélection par le Service des archives historiques) au cas d'un transfert.

Sur la base de ces dernières métadonnées, des opérations manuelles ou automatiques deviennent possibles, telles que la génération des rapports ou l'avertissement aux utilisateurs compétents de l'arrivée à échéance pour le transfert ou l'élimination des dossiers.

Cet exemple n'est qu'un aspect parmi d'autres de l'utilisation des métadonnées dans les systèmes de GED de la Commission européenne, mais vise à illustrer la logique de fonctionnement du système, tout en montrant que dans le monde du numérique, les métadonnées se prêtent à maintes possibilités en même temps qu'elles demandent une attention et une analyse très rigoureuses. À la Commission européenne, cette analyse est faite aussi bien au préalable (lors de la rédaction des Mda) qu'au fur et à mesure du développement opérationnel (en tant que partie intégrée du travail d'élaboration des spécifications des outils électroniques et de leur développement et mise en œuvre). Tout au long de ces analyses, les normes (formelles, comme par exemple ISO 15489, ou de facto comme par exemple les trois itérations du MoReq) ont souvent servi comme source d'inspiration et de référence. Pourtant, d'après notre expérience il ne nous a jamais été possible de reprendre ou d'utiliser ces normes telles quelles. Elles nous ont été très utiles, à maintes reprises, mais en les adaptant à notre situation spécifique et à nos besoins concrets.

---

<sup>26</sup>Cette date est calculée à partir de la date de clôture, en y ajoutant le nombre d'années de la DUA et en arrondissant à l'année entière qui suit. Pour un dossier avec la date de clôturé 1 mai 2013 et une DUA des 5 ans, la fin DUA est égale à 2019 et c'est à partir du début de cette dernière année que l'action post-DUA peut être effectuée.

#### 5.4.LA MISE EN OEUVRE DE ISO 23081 DANS LE CONTEXTE DU GOUVERNEMENT DU CANADA (NATHALIE VILLENEUVE - CANADA)

---

Au sein du Gouvernement Canadien (GC), la Direction du dirigeant principal de l'information (DDPI) du Secrétariat du Conseil du trésor (SCT) est responsable du développement de politiques régissant les domaines de la gestion de l'information et des technologies de l'information pour les institutions du GC. Le cadre de politiques en gestion de l'information du SCT inclut la Directive sur la tenue de documents, sous laquelle se trouve la *Norme sur les métadonnées*<sup>27</sup>. La première se concentre sur la gestion des ressources documentaires à valeur opérationnelle (RDVO) dans le contexte plus large de la gestion de l'information et la seconde fournit des indications précises sur la manière dont la gestion des RDVO est effectuée par les métadonnées. Les ressources documentaires (RD) comprennent toute « production documentaire sous forme publiée ou non, quelle que soit la source de communication, le format, le mode de production ou le support d'enregistrement<sup>28</sup> ». Les documents d'activité du GC sont l'un des types de RD, et les RD ont une valeur opérationnelle lorsqu'elles « permettent et documentent la prise de décisions à l'égard de programmes, de services et d'opérations continues, et appuient les exigences des ministères en matière d'établissement de rapports, de rendement et de responsabilisation »<sup>29</sup>. Les RD de valeur continue sont un sous-ensemble des RDVO.

La *Norme* sur les métadonnées indique les exigences pour trois domaines :

- la tenue de documents,
- la découverte des ressources Web,
- et les systèmes de gestion du contenu Web.

Les exigences pour les métadonnées pour la tenue de documents sont indiqués à l'Annexe B, et sont basés sur l'ensemble d'éléments génériques de métadonnées trouvés dans ISO 23081-2, Information et documentation - Processus de gestion des enregistrements -- Métadonnées pour les enregistrements - Partie 2: Concepts et mise en œuvre. Les éléments spécifiques inspirés de l'ensemble générique ISO sont identifiés dans cette annexe (s.3, s.4, s.5) et ils documentent certains aspects des entités de documents et des entités d'agents, aussi bien que de l'historique des événements d'un document. Ces éléments doivent être appliqués par les institutions du GC dans les dépôts de tenue de documents. Des métadonnées additionnelles peuvent être utilisées en fonction des exigences opérationnelles. Alors que techniquement les exigences de l'Annexe B pourraient être interprétées comme s'appliquant à tous les types de RDVO (pas uniquement aux documents d'activité), l'orientation centrée sur les documents d'activité est clairement définie dans la suite d'outils pour la tenue de documents dérivant de la *Norme* :

- l'Ensemble d'éléments de métadonnées pour la tenue de documents du gouvernement du Canada (RKMES),
- le Profil d'application de métadonnées pour la tenue de documents du gouvernement (RKMAP) du Canada,
- et la Ligne directrice relative à la mise en œuvre des métadonnées pour la tenue de documents du gouvernement du Canada (RKMIG).

---

<sup>27</sup><http://www.tbs-sct.gc.ca/im-gi/imp-pgi/imp-pgi-fra.asp>

<sup>28</sup> <http://www.tbs-sct.gc.ca/pol/doc-fra.aspx?id=18909&section=text>

<sup>29</sup>Ibid

Le RKMIG affirme que les RD telles que les publications et les bases de données ne sont pas comprises dans la portée<sup>30</sup>

Les exigences énoncées dans l'annexe B de la *Norme* représentent une réflexion ou mise en œuvre partielle d'ISO 23081-2, éléments de métadonnées génériques dans le modèle conceptuel des normes de métadonnées d'ISO. Ce ne sont pas tous les éléments de l'ensemble générique de métadonnées d'ISO 23081-2 qui sont présents, et de plus, seule une petite portion du modèle conceptuel y est décrit. Par exemple, seule une partie des niveaux hiérarchiques dans les classes d'entités de documents et d'agents est spécifiquement représentée et d'autres sont complètement absentes. Il faut ajouter que les relations entre les entités sont représentées de manière marginales et que les métadonnées de planification d'événements sont absentes. Cette décision a été motivée par la nécessité de diminuer la lourdeur et le coût d'exigences pour les exécutants, mais aussi parce qu'une approche graduelle de définition des métadonnées a été jugée préférable, afin de bénéficier de l'expérience des premières mises en œuvre et adaptations de la *Norme*.

En somme, même si la *Norme* sur les métadonnées est en droite ligne avec ISO 23081, elle ne comprend pas les particularités des éléments de métadonnées à mettre en œuvre, ni son inscription dans le cadre conceptuel d'ISO. Ceci limite la normalisation des métadonnées pour la tenue de documents d'ISO 23081 dans les dépôts du GC. Les carences entre la norme sur les métadonnées du SCT en lien avec la Directive sur la tenue de documents ont été identifiées et seront corrigées dans les révisions à venir. À titre d'exemple, les éléments spécifiques de métadonnées liés à la conservation et la disposition devront être ajoutés. Il y a aussi lieu de clarifier ou élargir comment ISO 23081 doit être réfléchi dans l'infrastructure de métadonnées pour la tenue de documents des institutions du GC dans les futures versions de la *Norme*.

## 5.5.POINT DE VUE DE SUR LE CLOUD STORAGE (JOHANNE. LESSARD - QUEBEC)

---

Dans un contexte de ressources partagées et très souvent externalisées, l'informatique nuagique<sup>31</sup>, prend une dimension particulière dès lors qu'il s'agit d'entreposer, voire de conserver des documents d'activité. Pour bien comprendre la portée de de l'informatique nuagique, il faut d'abord en comprendre les fondements pratiques.

- Les serveurs appartiennent à la compagnie de Services nuagiques,
- Les licences d'applications sont payées par la compagnie de Services nuagiques.
- La sécurité informatique est assurée par la compagnie de Services nuagiques.
- Les données résident sur les serveurs appartenant à la compagnie de Services nuagiques.
- Les serveurs sont localisés sur les propriétés de la compagnie de Services nuagiques.
- La maintenance des systèmes et applications est assurée par la compagnie de Services nuagiques,

---

<sup>30</sup>Final drafts available at [http://www.gcpedia.gc.ca/wiki/Metadata\\_Tools](http://www.gcpedia.gc.ca/wiki/Metadata_Tools) (internal GC site). Publication pending, 2013.

<sup>31</sup>« Informatique nuagique » et « Infonuagique » sont des traductions de *Cloud Storage* utilisée au Québec.

- Les ordinateurs appartiennent soit à l'organisation ou à la compagnie de Services nuagiques.
- Les données appartiennent à l'organisation.

L'entreposage nuagique possède un avantage auquel les gestionnaires sont très sensibles : l'avantage économique. En effet, ne plus avoir à acheter de serveurs, de licences de produits, ne plus avoir à assurer la maintenance et la sécurité informatique et/ou à faire l'achat et la gestion d'un parc informatique génèrent des économies substantiels pour une organisation, Plusieurs gestionnaires n'y regarderont pas à deux fois avant d'effectuer le changement de cap.

Pourtant, l'entreposage nuagique comporte aussi son lot de risques.

- Au niveau technique, l'organisation n'est plus maître des changements de version des applications qu'elle utilise, ce qui peut entraîner un manque d'intégrité des documents.
- Au niveau de la sécurité physique, l'organisation ne contrôle plus les accès des locaux où sont entreposées ses données, ce qui peut donner lieu à des accès non-autorisés aux données de l'organisation.
- Au niveau de l'accessibilité des données de l'entreprise, l'organisation n'a généralement aucun droit de regard sur les employés des Services nuagiques. Ceci peut donner lieu à un bris de confidentialité en regard avec des données sensibles.
- Au niveau de la sécurité des transmissions des données, l'organisation est à la merci du Service nuagique en cas de panne. Chaque heure de production perdue a un coup sur les produits/services de l'organisation.
- Au niveau de la localisation physique du Service nuagique, l'environnement légal du pays où il se trouve peut être différent de celui de l'organisation. Ceci peut donner lieu à des pratiques contraires à l'environnement légal dans lequel l'organisation opère et, en cas de litige (eDiscovery), il s'en suit un flou juridique. Même si les données appartiennent à l'organisation qui est par exemple située au Canada (ex. : non-soumis au PatriotAct), si les données sont entreposées dans un Service nuagique aux États-Unis (soumis au *PatriotAct*), on pourra facilement argumenter que les données se trouvant en sol américain sont effectivement soumises à la loi américaine pourtant inexistante dans le pays d'où origine les données. L'appareil législatif étant toujours à la remorque des avancées technologiques, aucun cadre juridique n'existe encore au niveau international à ce sujet.

Avant de passer la gestion de nos données à un Service nuagique, le gestionnaire de document et les Services juridiques de l'organisation devraient se pencher sur les items suivants :

- La protection et la sécurité de nos données;
- La continuation des affaires et des services;
- La sécurité des données (intégrité et fiabilité);
- Les modalités de récupération des données advenant une terminaison de contrat;
- Les problématiques de transmission de données inter-pays / états / régions / districts;
- Les modalités de transition (organisation/service ou service/organisation);

- Le niveau de service requis (*Service Level Agreement*) (disponibilité et récupération en cas de sinistre);
- La confidentialité;
- Les échelles de prix pour chacun des services offerts;
- La propriété intellectuelle;
- Les modalités de consentement à l'accès de données de l'organisation;
- Les modalités d'audit;
- Les taxes applicables à la transmission de données;
- Les assurances;
- La responsabilité et les dédommagements en cas de non-conformité au contrat établi; et
- La gouvernance (organisationnelle et en cas de litige).

Après analyse de tous ces risques, un « business case » positif d'impartition vers un Service nuagique est souvent plus difficile à prouver dépendamment de la situation juridique et du créneau d'affaires de l'organisation.

Les Services nuagiques ont donc développé la notion de Nuage privé (Privatecloud) afin de diminuer les irritants inhérents aux risques. Cette alternative ne diminue que partiellement l'ensemble des irritants soit la sécurité informatique, la propriété intellectuelle, la confidentialité mais des métadonnées de contrôles doivent être en place afin d'assurer la sécurité, l'intégrité, la fiabilité, la confidentialité et l'« auditabilité » des données. Il est donc recommandé que le système de gestion des documents d'activités employé dans un environnement nuagique soit certifié aux standards DoD 5015.2 ou MoReq2 puisque ces standards demandent des métadonnées d'audit sur les différents contrôles mentionnés plus haut.

## 6 CONCLUSION

---

Les révisions des normes ISO 23081 et ISO 15489 en articulation avec les normes de systèmes de gestion des documents d'activité ISO 30300 marquent un moment important dans l'évolution du corpus des normes de gestion des documents d'activité.

Tout d'abord, ces révisions donnent l'occasion de rappeler et de souligner la nécessité de disposer de métadonnées comme jeux d'informations à la fois pour décrire et pour gérer les documents d'activité et leur évolution.

Ensuite, si la preuve est faite de l'utilité des métadonnées, il n'en reste pas moins vrai que les normes existantes en matière de repérage, d'identification et de gestion des informations et des documents, en général dans les systèmes, ne répondent pas parfaitement aux besoins de gestion des documents d'activité.

En conséquence, un principe – ou une méthode - d'élaboration et d'évolution des métadonnées pour la gestion des documents d'activité est à concevoir afin de générer un modèle opérationnel adaptable aux différentes situations qu'il est possible de rencontrer dans les organisations. Dans cette perspective, et sur la base des normes ISO existantes et en particulier l'ISO 15836 dite du *Dublin Core*, il ne semble pas que le travail de réflexion et de formalisation soit compliqué à conduire et à réaliser. Il devrait rapidement aboutir à définir un modèle constitué d'un noyau de métadonnées et de couches spécifiques dotées de règles d'extension permettant de répondre à la diversité des documents d'activité et à la diversité de leur gestion existantes dans les organisations.

En revanche, la question qui se pose maintenant est celle du positionnement et de l'intégration de ces développements dans le corpus des normes et des rapports techniques connexes à l'ISO 30300. Faut-il largement adapter la norme ISO 15489 du *records management* très bien comprise et largement mise en œuvre par les professionnels ou bien revoir la norme ISO 23081 dédiée aux métadonnées ? La pression économique entraînée par le marché avec le *Cloud Storage* et le *Big Data* ne fait qu'amplifier le besoin et accélère l'obsolescence des référentiels normatifs existants.

Enfin, l'enjeu est aussi dans la collaboration active des pays francophones oeuvrant dans le travail normatif international afin de poursuivre l'effort de mise en commun de la langue française.

Dans ce contexte nous espérons que ce Livre blanc contribuera aux travaux sur les métadonnées à mener dans le cadre, des comités et groupes de travail du SC11 ainsi que lors des réunions de l'ISO qui se tiendront à Paris au cours du mois de juin 2013.

## 8 ENGLISH VERSION

### 9 FOREWORD: THE REASON FOR THIS WHITE PAPER

---

A large-scale standardisation project is currently under way concerning records management. A series of White Papers published by the French-speaking community aims to contribute to this project.

**White Paper no.1.** This was presented at the TC46/SC11 meeting held in Sydney in 2010 and enabled the introduction of the ISO 3030X series of management system standards. It laid out the necessary changes to ISO 15489 on records management. This initial white paper enabled the start of a French-speaking approach embodied by the French version of ISO 30300 and 30301, which was approved at the International Organisation for Standardization (ISO). It introduced and asserted the role of the CN11 as a proactive entity.

**White Paper no.2.** This was presented at the TC46/SC11 meeting held in London in 2011. Its influence was a major factor in the revision of ISO 15489. The standardisation team from Canada wrote a summary in English.

**Bilingual White Paper no.3.** The theme of this White Paper was models and diagrams representing records management systems. It was presented at the TC46/SC11 meeting held in Berlin. This version, which was fully translated into English, received numerous contributions from around the world. At the same time, in France, it enabled several processes of reflection to be undertaken, particularly concerning the relevance of the three ages paradigm in archiving in terms of the digital representation of records.

**Bilingual White Paper no.4** is to approach the issue of metadata and their standardisation in order to prepare for the TC46/SC11 meeting taking place in Paris in June 2013. This is a particularly sensitive area, as "data about data" allow us to identify, find and exchange information between systems. What's more, some metadata can help to provide evidential value. As a result, and with the arrival of big data and the economic activity connected with cloud computing, the management and prior definition of metadata are therefore strategic in nature.

The standard ISO 23081: Metadata for records from January 2006 defines the principles that govern metadata linked to records and their management. In its introduction, the standard states: "ISO 23081 establishes a framework for creating, managing and utilising metadata associated with records management and explains the principles that govern this data."

The fact that the process of systematically revising ISO 23081 is taking place in 2013, as is the revision of ISO 15489, constitutes an exceptional opportunity for new reflections to take place.

This White Paper aims to form part of this context by attempting to begin to answer various questions.

#### DO RECORDS MANAGEMENT METADATA HAVE A SPECIFIC CHARACTERISTIC?

---

Numerous standards already exist concerning metadata and more broadly mechanisms for locating information and documents within systems. ISO 15836 on so-called Dublin Core metadata or the exchange of information, as well as Z39.50 and ISO 23950, cannot be ignored. The question is whether records management requires specific metadata in other words, a set of metadata without which the ISO 30300 series of standards and the associated technical standards, particularly ISO 23081 and ISO 15489, could not be implemented reliably, securely and over the long term.

This White Paper will address reflections currently under way concerning the creation of an ISBN<sup>32</sup>-style unique identifier for records or a persistent indicator (see DIN 31646).

In this context of change in the body of standards on records management, two main working hypotheses are presented with the aim of optimising questions regarding metadata:

- Revise the ISO 23081 standard to obtain a more accurate definition of the set of metadata so as to ultimately enable smoother implementation of the standard.
- Integrate into the revision of the standard ISO 15489 elements concerning the requirements related to the set of metadata, such as for example creating a section 3 specific to metadata.

This White Paper will attempt to shed new light on these questions and establish a formal viewpoint.

#### HOW SHOULD WE ENVISAGE THE EVOLUTION OF METADATA IN VIEW OF THE LIFECYCLE?

---

Records form part of a lifecycle marked by business events to which they are linked, as well as technological developments that arise, such as changes in physical media for example. When these management events occur, it is necessary to evolve the metadata without compromising the integrity of the document or the reversibility of the process. Therefore, over time and as a result of successive changes in the medium, the "records management" pairing becomes a Russian doll-style casing system that can be expressed as (((document-metadata)-metadata)-metadata), etc.

The standard ISO 23081 defines this need as follows: "All metadata about the record and those accruing in its management and use also form a record: the metadata document that also has to be managed."

As a result, the lifecycle/metadata pairing is very specific to records management and will also be commented on in this White Paper.

Lastly, the expression cloud computing<sup>33</sup> recently emerged as a result of major commercial efforts and service offerings from providers. There are various offshoots,

---

<sup>32</sup>International Standard Book Number.

including cloud storage, which is of key importance as far as records management is concerned. The members of the working group wanted to make several comments on this matter.

This White Paper contains three main sections that analyse and attempt to answer the aforementioned questions:

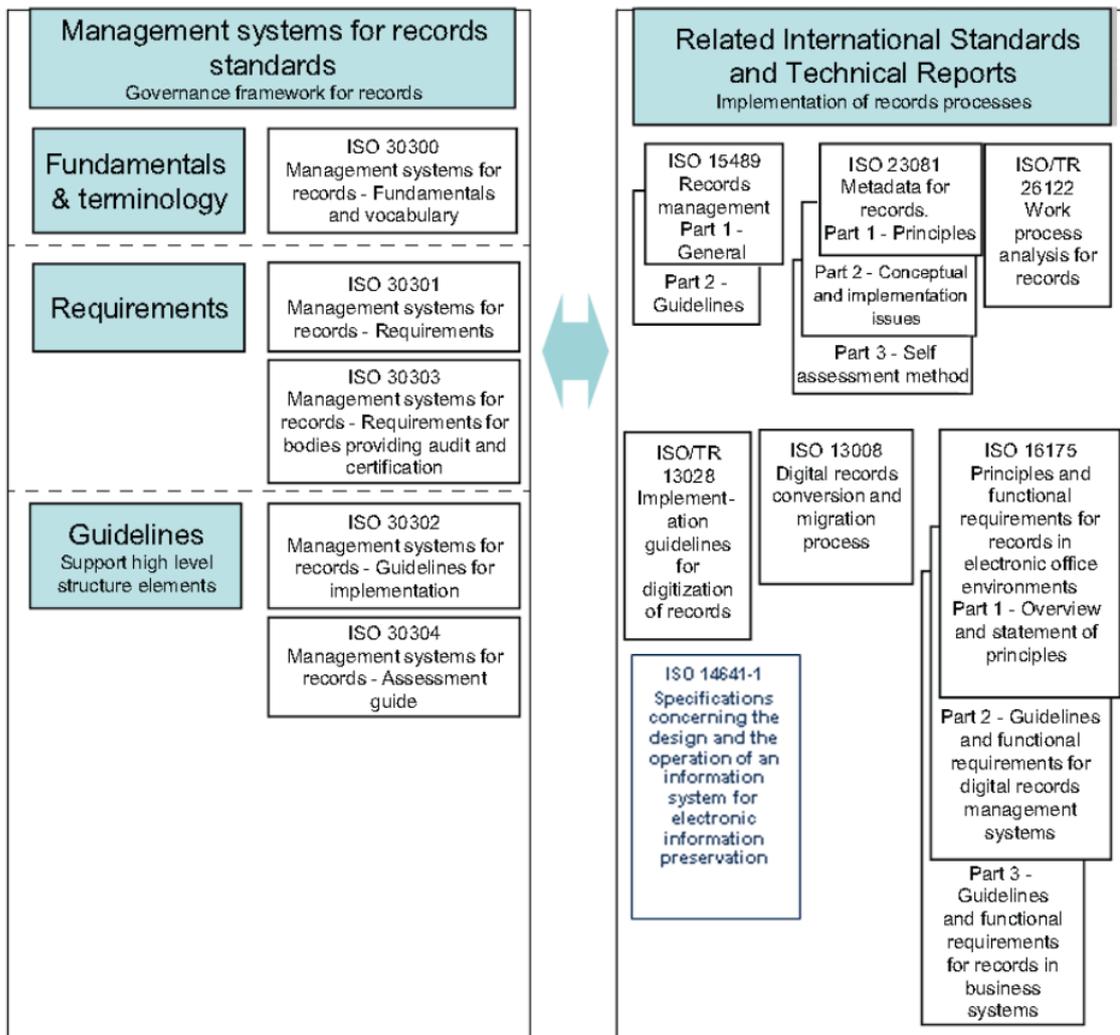
- Section 3 contains definitions and provides a reminder of what metadata are.
- Section 4 states requirements concerning records management metadata.
- Section 5 contains case studies.

---

<sup>33</sup> French expressions exist, such as "informatique nuagique" and "infonuagique". It will be important to take this terminology into account during upcoming translation work in keeping with the work carried out on the 3030x series.

## 10 THE CURRENT STANDARDS CONTEXT

The standards context currently being developed can be summed up by the following diagram published in the standards ISO 30300 and 30301. The meeting being held in Paris in June 2013 will contribute developments to the ISO TC46/SC11 projects.



## 11 METADATA, REMINDERS AND EXISTING STANDARDS

---

ISO 30300 states in section 3.1 "Terms concerning records" that metadata are "data describing the context, content and structure of records as well as their management over time".

ISO 30301 states in section 8.1 "Concepts of processes linked to records" that when creating an RMS<sup>34</sup> it is necessary to "determine the content, context and control information (metadata) that must be included in records".

### 11.1. DEFINITIONS, REMINDERS OF FUNCTIONS AND SPECIFIC CHARACTERISTICS OF METADATA

---

#### 11.1.1. DEFINITION

---

Metadata (literally, "data about data") is structured information that allows us to identify and describe the documents and/or data produced in connection with an activity and to reuse them in the short, medium and long term in accordance with users' requirements. As such, they are essential for presenting document production from a technical as well as intellectual viewpoint. They therefore enable information to be made accessible and usable. In this sense, metadata are an essential tool in records management.

Metadata can be likened to a label on a tin. This label is essential for providing information to the consumer: it gives a set of indications about the tin's contents (the brand, the weight, details on the ingredients and their energy value, the date of packaging, the expiry date, the name of the producer/distributor, its address, etc.), thereby enabling essential information to be provided. A document's "label" works in the same way: it enables its content to be easily identified in its production context in relation to the work process (or activity) thanks to structured information such as the creation date, the creating entity, etc. By documenting the information produced, the metadata brings added value to identification of the document. In the same way that it would be difficult to identify a tin containing peas in a batch of unlabelled tins, it would be difficult to identify a document concerning an annual report for a given year among a series of documents that have no metadata.

In his article "Histoire(s) de notices"<sup>35</sup> (A History of Records"), Gautier Poupeau says that "the use of the term metadata constitutes a semantic change whose scope should not be ignored. This term embodies the fact that observers are no longer just interested in a record [a descriptive record accompanying a document for example] as a whole, but in the various portions of information that it contains through the notion of "metadata". As such, information is grasped through that which characterises it at the most detailed level in other words, there can be no descriptive record without record fields.

---

<sup>34</sup> Records Management System

<sup>35</sup> POUPEAU (Gautier), "Histoire(s) de notices" ("A History of Records") in: *The digital document in the era of the web of data*, INRIA seminar (Carnac, 1-5 October 2012), Paris: ADBS, 2012.

---

### 11.1.2. METADATA PERFORM VARIOUS FUNCTIONS

---

All digital documents must be accompanied by metadata. Several categories of metadata corresponding to different functions and at varying levels of requirement can be identified. In records management, the following categories can be identified:

- Technical metadata
- Administrative metadata
- Descriptive or content-based metadata
- Records management metadata, which can include technical, descriptive and administrative metadata at the same time.

---

#### TECHNICAL METADATA

---

The first layer of metadata is mainly comprised of technical information that a machine needs in order to reinterpret the files. The distinctive feature of these metadata is that they are automatically generated by the information systems that were used to produce the documents/data. This minimal information could be likened to the basic data written on an envelope (first name, last name, street number, town, post code), without which it would be difficult for a postman to deliver mail to its recipient.

However, these technical metadata are insufficient from a records management viewpoint. We might consider, for example, that an electronic records management system that aims to guarantee the authenticity and integrity of documents should integrate other specific metadata such as timestamping, a calculation of the footprint of documents and the MIME type of files verified by specific tools such as Jhove and DROID (Digital Record Object Identification). It is therefore possible to identify a more comprehensive set of technical metadata depending on the management objectives that have been defined.

---

#### UNIQUE IDENTIFIERS

---

Often likened to technical metadata, unique identifiers can be considered a separate typology of metadata given the specific nature of their role. The identifier serves as a diacritical criterion that allows a resource to be identified with certainty relative to an activity process or within an information system.

Unique identifiers are based on standardisation rules, which should not be confused with simple file naming rules.

---

#### IDENTIFYING A RESOURCE RELATIVE TO AN ACTIVITY PROCESS

---

##### **The example of the NOR number**

The NOR is a unique identifier provided by the S.O.L.O.N application to identify legislative and regulatory bills due to be published in the Official Journal of the French Republic. It is made up of 12 alphanumerical characters forming five cohesive sets:

E.g. MAEH1252486D

- A first set made up of three letters corresponding to the ministry issuing the text (in reference to a table updated upon every government changeover).
- A second set made up of one letter corresponding to the department of the ministry issuing the text (in reference to a table updated upon every government changeover).
- A third set made up of two letters corresponding to the year the text was created (this figure is automatically provided by the system and updated every year).
- A fourth set made up of five figures corresponding to a numerical series

incremented by the system generating the text whenever a text is created (the numerical series is restarted every year).

- A fifth set made up of one letter corresponding to the type of text created (in reference to a table).

### **The example of the ECLI number<sup>36</sup>**

The ECLI number (European Case Law Identifier) is an identifier designed to aid the correct and unequivocal quotation of European Union-related legal rulings given by European and national courts.

The ECLI identifier is made up of five elements:

6. ECLI: An abbreviation indicating the nature of the identifier
7. A country code or an organisation code: for France this is "FR"
8. The abbreviation of the court
9. The year of the ruling
10. An order number

Each element is separated by a colon ":" »

---

### ADMINISTRATIVE METADATA

There is also a category of metadata that is used to administrate a system, i.e. manage users' profiles and their rights.

---

### DESCRIPTIVE METADATA OR CONTENT-BASED METADATA

The role of these metadata is to describe the content of documents in order to facilitate access to them. This is undoubtedly the most common category of metadata.

The standardisation effort in this area is significant, examples being thesauruses and other forms of controlled language. In the field of descriptive indexing, there is no single model: by definition, each indexing language is created to cater to a specific field of activity.

There are a few specific applications:

**EUROVOC<sup>37</sup>**: a multilingual thesaurus used by the European Parliament, the Publications Office of the European Union, national and regional parliaments in Europe, national administrations and certain European organisations. It covers all the areas of activity of the European Union and enables documents to be indexed in the document systems of European institutions and their users.

**RAMEAU**: used to index bibliographic content in libraries.

**Thésaurus W**: used to describe the collections in the archive departments of local authorities in France.

---

<sup>36</sup> <http://www.parcesmotifs.net/spip.php?article10>

<sup>37</sup> [http://europa.eu/eurovoc/sg/sga\\_doc/eurovoc\\_dif!SERVEUR/menu!prod!MENU?langue=FR](http://europa.eu/eurovoc/sg/sga_doc/eurovoc_dif!SERVEUR/menu!prod!MENU?langue=FR)

Although all of the aforementioned categories are fundamental, they are insufficient for covering the scope of records management.

---

### 11.1.3. THE SPECIFIC CHARACTER OF RECORDS MANAGEMENT METADATA

---

Records management metadata are necessary for retaining and presenting a document in its production context and thereby preserving its link with the work process (or the activity). They also enable management of the lifecycle of information (the retention periods or retention timetable, the ultimate fate). They help to ensure the authentic nature of documents (evidence management).

The standard ISO 23081 describes a general conceptual framework used to identify the main functions of metadata for records management:

- Protection of the probative nature of documents (this point is addressed in more detail in 3.2.2)
- Guaranteeing their accessibility/usability over time
- Guaranteeing the authenticity, reliability and integrity of documents over time
- Managing access, rights and protection (of personal data)
- Interoperability
- Maintaining the links between documents and their creation context through a reliable structure
- Identifying the technical environment of document creation
- Facilitating the migration of documents from one technical environment to another

---

### THE CHALLENGE OF INTEROPERABILITY: THE IMPORTANCE OF STANDARDISATION

---

Metadata are also a response to growing needs in the exchange of information. In this sense, interoperability has become a key issue in improving the management of information and its dissemination.

Standardisation for interoperability purposes was practised very early on by libraries, with the codification of the description of their documents. This resulted in the structuring of documents in the form of paper records containing various fields, then the definition in 1954 of a standard, the ISBD (International Standard Bibliographic Description), which aims to establish consistency when sharing bibliographic information. To do so, the ISBD "determines which elements of data need to be noted or transcribed into a specific sequence that serves as the basis for describing the resource being catalogued"<sup>38</sup>. With the widespread availability of catalogues online, new standards concerning record fields subsequently emerged (MARC formats).

The standardisation of the description of collections in archive services came much later: it wasn't until 1994 that the ISAD(G) (General International Standard Archival Description) standard was published and a stable version of the DTD EAD (encoded archival description structured with XML language) didn't emerge until 2002. The EAD is currently being revised to move the DTD towards an XML setup that will facilitate interoperability with other formats of metadata.

---

<sup>38</sup> [http://www.bnf.fr/documents/isbd\\_integre.pdf](http://www.bnf.fr/documents/isbd_integre.pdf)

Several attempts to model metadata have been made since the 1990s<sup>39</sup>:

- 1996: the CIDOC-CRM (*Committee on Documentation – Conceptual Reference Model*): this provides a conceptual reference framework to describe all of an object's interactions with other entities in an object-oriented model.
- The FRBR (*Functional Requirements for Bibliographic Records*), published in 1998: these define three groups of entities to describe documents, physical persons and legal entities, and the various concepts that may be the subject of projects.
- The *Dublin Core*, which in 2007 defined an abstract model (DCAM, *Dublin Core Abstract Model*).

The development of metadata for records management came even later. The standard ISO 23081, presents a conceptual framework for managing these metadata, was published between 2004 and 2006. Standards such as MoReq and ISO 16 175 (ICA-Req), aims to provide guidelines along with practical models for records management, subsequently added to the body of standards, although there was no unified and standardised model for metadata.

However, the standard does not recommend a specific metadata model and does not provide a checklist. At present, there is no French-language model in terms of minimum requirements and a homogenous vocabulary.

What is needed is a model that can converge the various scattered existing models as well as existing standards, including ISO 15489, ISO 30300, ISO 23081 and ISO 16175 (ICA-Req). It is also necessary to bear in mind that there cannot be a common minimum foundation given that each activity process is by definition unique.

This challenge is becoming a key issue in public administration: with the effort to streamline public finances conducted through recent reforms (LOLF<sup>40</sup>, RGPP<sup>41</sup>, modernisation of public action), the need for more transparency, traceability and interoperability (e.g. interoperability of systems to facilitate the definitive electronic archiving of metadata) has become more apparent.

## 11.2. EXISTING STANDARDS, THEIR LIMITATIONS AND THEIR RELEVANCE

---

There is a huge number of sets of metadata. There is virtually one for each specific example of data: to describe formats, quantify volumes, evaluate exchanges and flows, identify the "originals" of "copies", trace the circulation of data in the producer entity, etc. However, strangely, sets of metadata specific to records management have never been stabilised in terms of their form. Recommendations exist but these are often unsuitable, as they are too generic.

---

<sup>39</sup>See Gautier Poupeau

<sup>40</sup>LOLF: the French acronym for the organic law concerning finance laws.

<sup>41</sup>RGPP: the French acronym for the general revision of public policies.

---

### 11.2.1. POSITIONING OF THE STANDARDS ISO 23081, ISO 15489 AND 30300

---

Metadata play a decisive role in the process of monitoring and controlling the lifecycle of records and in the traceability of actions in, around and about data and records: records management standards place a particular emphasis on the description of content and the structures of sets of metadata. "Records management metadata can therefore be used to identify, authenticate and contextualise records and the people, processes and systems that create, arrange, maintain and use them and the policies that govern them [...] metadata ensure authenticity, reliability, usability and integrity over time and enable the management and understanding of information objects", standard ISO 23081-1 (section 4).

The standard ISO 15489-1 defines the metadata specific to records management as data describing the context, content and structure of records and their management over time (ISO 15489-1 (section 3.2), which is repeated in the standards ISO 23081-1 (section 4), ISO 30300 and 30301 and expanded upon in the standard ISO 16171 (4312)). This definition is then repeated in the standard ISO 23081-1 (section 5.2.1), which distinguishes two types of metadata:

- Metadata concerning context, content, appearance and structure used to create or capture records.
- Metadata for the management of records and processes in which the document is used, including changes made to the content, structure and appearance.

The method for managing these metadata must be consistent over time, regardless of the environment (ISO 15489-1, 16171 and 23081-1).

All of the standards mentioned stress to various degrees of detail the definition of the set of metadata in the process of implementing a records management policy (ISO 15489 section 7.1 and ISO 23081-1 (section 5.2.1), the structuring of the set (ISO 15489-1 section 9) and the management of the set (ISO 23081-1, section 5.2.3 and ISO 16171 terms and definitions, and ISO 15489-2 (section 4.3.1.)).

However, although these four standards complement each other relatively well in an attempt to define requirements in terms of metadata, they do not recommend a standard set suited to records or a grid for assessing existing sets of metadata in records management systems.

---

### 11.2.2. ISO 23081-1 AND THE NOTION OF EVIDENTIAL VALUE

---

At the beginning of 2012, the translation of ISO 23081-1 - Information and Documentation – Records management processes – Metadata for records – part 1: Principles was submitted for a public enquiry by AFNOR. Because the content of this standard could not be revised, as it had already been published at the ISO in English in 2006, the purpose of the enquiry was to determine whether the French translation could be used in the form submitted for the enquiry.

The main objection raised involved what was described as a translation error, which upon closer inspection was a result of the ambiguity of the English version. The translation in question was in the section "5.1 Purposes and benefits of records management

metadata<sup>42</sup>". The objection centres on the fact that the text "indicates that it is sufficient for 'records' to be accompanied by metadata for them to have 'evidential value'. This creates a very troublesome confusion between the notion of 'necessary condition' and 'sufficient condition'." In other words: not all metadata have evidential value, only some, and in all cases metadata alone are not sufficient to guarantee this value.

In concrete terms, the translation initially proposed and the ambiguity of the English text made and continue to make this text difficult to use or even unusable in its current state in connection with commercial activities related to electronic archiving, as it is incompatible with the reference standard in this field, NF Z42-013: 2009 (and its international equivalent ISO 14641-1: 2012).

From a legal viewpoint, it is also possible to question the relevance and solidity of the standard ISO 23081-1 as published at present. Although "certain" metadata could confer evidential value, the standard does not aid the specific construction of a reference set of metadata: the recommendations are general and do not contribute much.

A French consensus was reached during the study meeting held in July 2012, by stating that "certain metadata help to establish" probative value. However, aside from this "prop", the commentators present at the meeting strongly pushed for a revision of the initial standard, for the reasons stated.

---

### 11.2.3. THE STANDARD ISO 15836, OR DUBLIN CORE

---

A set of metadata that might be deemed suitable for records management would be the *Dublin Core*: this is the subject of the standards ISO 15836 and ANSI/NISO Z39.50, is based on the standards of the W3C and is widely used by information management systems in all fields and not just those of libraries. Dublin Core is a set of metadata primarily aimed at document research and is based on 15 metadata elements, which can be qualified (i.e. the definition of the elements is restricted) or not. It is also worth noting that the standard ISO 1583 only aims to describe metadata and gives them a French translation, with no implementation method. The main goal of Dublin Core is document research and this goal is only one of the aspects of a records management policy.

Paradoxically, the extremely widespread use of this set of metadata by an impressive variety of organisations should be cause for concern and lead us to question whether there is a significant margin of free interpretation given to each organisation to fill the content of each metadata item. For example, the content of the metadata item "date" could be completed by one organisation based on a document's creation date, whereas another organisation might choose to only indicate the last date modified. The absence of a clear timestamping concept or a unique identifier is a source of risk in the implementation of systems.

As such, for a set of records management metadata and following the guidelines development in the standards 15489 and 30300, the 15 items constituting Dublin Core are not an adverse factor, but they are insufficient. Dublin Core cannot be applied "as is" to records management and should be complemented and adapted.

---

<sup>42</sup> "Metadata supports business and records management processes by protecting records as evidence", translated as "les métadonnées (...) confèrent aux documents leur valeur de preuve" .

---

#### 11.2.4. FOCUS ON THE UNIQUE IDENTIFIER

---

One element that is particularly representative of the reflection process prior to the creation of a set of metadata concerns the identifier of a record. Although in some companies or organisations, a nomenclature for naming documents has been put in place as part of a records management procedure, the electronic context concerning metadata is another matter entirely.

Part 9 of the standard ISO 15489-1 states that the identifier is an essential element of descriptive metadata. It must be unique and enables a document to be referenced in an organisation's classification plan. The standard ISO 15489-2 states that the identifier is generated when the record is integrated into the management system.

There are several solutions for achieving compliance in these standards requirements, including "house" naming systems and the use of identifier systems, such as the ISSN and DOI (Digital Object Identifier), developed by third parties, called "naming authorities".

## 12 REVERSIBILITY, SECURITY, CLOUD: WHICH REQUIREMENTS IMPACT METADATA?

---

---

### 12.1. REVERSIBILITY REQUIREMENT

---

Reversibility is a process that enables the retrieval of data from one information system for transferral to another system. This operation is also called migration. It is necessary when replacing a system or changing a service provider that safeguards data.

Backup is different from migration, as its main aim is to restore a system identically following an incident that damages the system's data.

In the records field, migration concerns digital objects (documents, files, photos, etc.) and the metadata that characterise these objects.

Three elements are necessary in order for a system to be reversible:

- The source must have a function that allows the data to be selected and exported; for example, documents accompanied by their metadata.
- The target must have a function that allows the data to be imported and utilised.
- A transfer format compatible with the source and the target must be specified.

Except if the source system and the target system come from the same vendor handling data migration, it is necessary to choose and configure standardised formats such as ISO 22938 (Document management - Electronic content/document management data interchange format) and ISO 2709 (Exchange format for bibliographic data).

Whatever the type of application, the existence of operational reversibility is a key element in the longevity and security of data held over time in a system.

## 12.2. SECURITY REQUIREMENT

---

The security of RMS is of course linked to that of the metadata associated with records. This is why this White Paper aims to place a special emphasis on the metadata-security pairing in view of the most recent security-related standards.

Metadata could provide elements relating to sensitivity and the security requirements associated with information. This could facilitate the processing of this information, concerning security, throughout its lifecycle.

The necessary elements stem from an analysis of the risk associated with this information. They may concern the standard criteria for the analysis of risk in information security (availability, integrity, confidentiality), as well as those addressed in a complementary manner during this analysis (assignability, authenticity, longevity, etc.). More generally, they may also refer to the level of sensitivity of information as it will be defined at the end of the analysis.

The JTC1 is, at the ISO, the Technical Committee that manages information technologies. Within the JTC1, the SC27 is a sub-committee dedicated to security. The SC27 is comprised of five working groups, including WG1 and WG4, which are in charge of developing standards in the 27000 series built around the notion of ISMS (Information Security Management System).

The three central standards in this series are:

- ISO/IEC 27001, a requirement standard that enables the analysis of compliance and certification in information security.
- ISO/IEC 27002, a best practice standard that is associated with the previous standard and encompasses 130 measures.
- ISO/IEC 27005, a standard on managing information security risk that provides methodological support for the previous two standards.

The 27000 series is currently made up of just over 30 standards.

The most recent developments are as follows:

The standards below represent, firstly, the recent developments in the 27000 series and, secondly, current developments concerning cloud computing. The 27000 series currently comprises around 30 standards in the context of the certifying standard ISO 27001.

- ISO/IEC 27035 - Information security incident management – August 2011. This is a bottom-up standard that concerns information security incident management.
- ISO/IEC 27037 - Guidelines for identification, collection, acquisition and preservation of digital evidence – October 2012

This standard aims to guarantee the integrity of digital evidence in order to authorise its use in legal frameworks. It introduces additional criteria concerning the precautions to take in data management, notably: auditability, repeatability, reproducibility and justifiability.

### ISO/IEC 270XX STANDARDS LINKED TO THE CLOUD

---

Three standards concerning cloud computing are currently being prepared at the SC27:

- ISO/IEC 27017 - *Guidelines on information security controls for the use of cloud computing services based on ISO/IEC 27002 – WD 3*

This standard aims to allow the use of the standards 27001/27002 to be expanded to cloud computing.

- ISO/IEC 27018 - *Code of practice for data protection controls for public cloud computing services – WD2.*

This standard is developed at the WG5 (*Privacy technologies and Identity Management*). It concerns the protection of data and more specifically personal information.

- ISO/CEI 27036 – *Part 5: Information security for supplier relationships - WD*

This project is currently in five parts (*Overview and Concepts, Common Requirements, Guidelines for ICT Supply Chain security, Outsourcing, Guidelines for security of cloud services*). The fifth part, which concerns the security of services in cloud computing, is in WD.

### 12.3. COMMENTS REGARDING CLOUD STORAGE

---

The expression cloud computing emerged a few years ago as a concept in scientific literature and as a commercial offering from IT service providers. Seen as a new promised land, this expression has often been used without much consideration.

The NIST (National Institute of Standards and Technology) offers the following definition: "Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."

In this context of shared and very often outsourced resources, cloud storage acquires special importance when it comes to storing or preserving records. Although cloud storage has a number of benefits, particularly economic, it also has a number of aspects that require particular vigilance, such as integrity, security, reversibility, portability, resilience and continuity of service.

As a result, two comments can be made. The first is that cloud storage is a technical solution. It is not a product of functional considerations, particularly those concerning metadata. However, the second comment is that it is essential for all organisations considering cloud storage services to be attentive to the contractual aspects that link them to a supplier in terms of security, longevity, data protection and reversibility.

## 13 CASE STUDIES AND VIEWPOINTS

---

---

### 13.1. LINKLATERS, INTERNATIONAL LAW FIRM (SYLVIE FORASTIER - LUXEMBOURG)

---

#### 13.1.1. OBJECTIVE: TO ANSWER THE MAIN QUESTIONS ABOUT FILE IDENTIFICATION AND LOCATION.

---

To do so, data need to be distinguished from metadata. Data are something that will be of immediate use to a user in identifying a file and locating it (e.g.: name of the file, physical form, the number of occurrences, location, status, etc.). Depending on the nature of the metadata, they will not be used during every search but will serve to refine or clarify a

search when that search becomes complex. They will also be used to ensure traceability throughout the lifecycle of a given file. In a universe in which the file is multiform and is comprised of paper documents, electronic documents, emails and databases, metadata will be of the following nature:

---

### 13.1.2. METADATA ABOUT PAPER/ELECTRONIC DOCUMENTS.

---

These are mostly created when a document is produced. Some are automatically generated, such as the document number, the creation date, etc. Others are manual (such as the reference of the file). They are primarily intended for use by the public and the administrator, although each uses them differently. Administrators identify the source of the document using its references (for example: the date, the creator's initials, the number of the file, etc.) and define the destination of the document in order to provide input for the appropriate file. Meanwhile, users (which may not be the creator) use these same metadata for search purposes. They verify that users have achieved their goal and have found the document required.

These metadata are codified uniformly within a company (positioning on the document, abbreviations, etc.), which makes managing them fairly easy. However, the utilisation of these metadata relies on their accuracy, particularly when it comes to metadata that are manually encoded.

---

### 13.1.3. METADATA ABOUT PHYSICAL FILES.

---

These are generated by the database (see standard metadata - in the database), although the operator selects the most significant among them, as not all can be placed on the physical file (risk of confusion: only metadata such as the barcode and the archive number are relevant to state on the physical file). Instead of the excluded metadata, the administrator will add others that convey meaning (a colour code for example). As with metadata about documents, the use made of these varies for administrators and users, i.e. administrators will be aiming to make the files available while users will be searching for the files.

These metadata are codified by the administrator: therefore, the metadata that are the most useful to users must be determined. This stage may take time, as it requires knowledge of business processes. The check made on these metadata must be rigorous and clearly communicated, as they are liable to evolve (for example, an addition to the colour code) and they convey information: the incorrect usage of these metadata could create incorrect information.

---

### 13.1.4. METADATA FROM THE DATABASE.

---

#### STANDARD METADATA

---

These metadata are standard metadata, as they will be present regardless of the tool.

They are created by the tool itself and neither the administrator nor the public can influence them (barcode, archive number, creation date, archiving date, migration information, record of loans, etc.). What's more, the use that users can make of these metadata is very restricted and because access to certain metadata is not possible for users, they will have to contact the administrator if required. These metadata are mostly

for use by administrators, who will decide which are available for users (see metadata about physical files) and use this information for general administration of the records management department.

These metadata are the easiest to manage, as they are codified and created by the tool itself, without any risk of error.

#### PERSONALISED METADATA

---

In addition to the aforementioned metadata categories, the database contains personalised metadata that are linked to a company's specific needs. Created by the administrator, they are generated on a case-by-case basis rather than in a standardised way and must be accessible by users when they feel they require them (example: accounting information).

There are also personalised metadata for administrative use by the records management department (example: archiving operator agent).

These metadata are codified by the administrator: based on knowledge of the business processes, it will have been necessary to previously determine the most useful metadata for users, decide on a codification and then code them so that users can access them.

#### COMMUNICATION METADATA

---

These are created by users (e.g.: email, deposit slip, etc.) and are intended for the administrator. These metadata serve a purely administrative purpose and are used for general management of the records management department (creation request, archiving order, etc.). The main objective is to ensure the traceability of the instructions given.

They are very difficult to manage, as they have multiple sources and exist in great numbers. As a result, they are difficult to utilise (finding an email amidst 5,000 archiving orders). One solution proposed is to compile this information so that it is available in a single source. For example, compiling archiving orders in an Excel file enables the instructions given to be found using the "filter" function and allows users to locate the original instruction more easily and quickly.

#### MIGRATION METADATA

---

In the event of migration from one database to another, metadata give administrators the ability to verify the source and nature of information prior to the migration. The aim is to preserve a copy of the original data (before migration) so as to ensure that the migration takes place correctly. The objective of preserving these metadata is for use by the administrator in the general management of the records management department.

The variety of forms and sources of metadata require various management methods. There is not necessarily a link between all metadata. Nevertheless, all of these metadata coexist for a given file. In addition, metadata must be entered in fields that either the administrator, the user or both can query. The importance of metadata should not be underestimated, as they form the foundation of any successful data management system.

## 13.2. VIEWPOINT ON METADATA IN ENTERPRISES (LUCAS COLET - LUXEMBOURG)

---

Within the framework of business usage, the metadata reference system of the OAIIS fairly naturally emerges as an initial solution for constructing a set of metadata and classifying them under categories. Nevertheless, in practice it appears that three categories are predominant:

- Technical metadata (often called "management" metadata), encompasses all metadata concerning the longevity of information, its medium, its end purpose, and its use and access, which will therefore be primarily used by the manager (software or otherwise) of information storage. They may also be used for creating an evidential environment. These technical metadata may draw on MoReq2, which provides a relatively comprehensive overview of them.
- Business metadata (often called "descriptive" metadata), are specific to the core business and facilitates identification and searches in a given context (an HR department will have different metadata at this level to the Billing department, for example).
- Other metadata, which facilitate the presentation of information, searches and anything that is not covered by the other two categories.

Nevertheless, when constructing a set of metadata it is best to draw on a standard so as to enable maximum interoperability between these metadata.

### 13.3. METADATA AND THEIR ROLE IN ELECTRONIC DOCUMENT MANAGEMENT – AN OVERVIEW OF THE EXPERIENCE AT THE EUROPEAN COMMISSION (MATTIAS HALLIN)

---

In 2002 The European Commission began the process of modernising its document<sup>43</sup> management and archiving, in order to increase the efficiency and transparency of its work and to ensure that at any time, it can give account for matters for which it is responsible.

This involved, inter alia, the adoption by the Commission of two decisions annexing to its Rules of Procedure provisions on document management<sup>44</sup> and on electronic and digitised documents<sup>45</sup>. These two decisions were completed between 2003 and 2008 by Implementing Rules (IR, recast in 2009 as a single coherent text<sup>46</sup>), and by a common retention list (CRL, revised in 2012)<sup>47</sup>.

Those rules include a list of some 50 metadata (mandatory or voluntary)<sup>48</sup> covering the management of the entire life cycle of documents, files and the filing plan of the Commission. The list is sub-divided into six groups: registration of documents<sup>49</sup>, headings of the filing plan<sup>50</sup>, files<sup>51</sup>, preservation of files<sup>52</sup>, transfer<sup>53</sup> and elimination<sup>54</sup>.

---

<sup>43</sup> In this text we don't use the term "record", only "document", in order to respect the terminology of the legislative framework of the Commission (where the latter is the only term used, both in English and French, although in practice its meaning corresponds very closely to that of "record").

<sup>44</sup> Decision 2002/47/EC, ECSC, Euratom of 23 January 2002, OJ L 21, 24.1.2002, p. 23.

<sup>45</sup> Decision 2004/563/EC, Euratom of 7 July 2004, OJ L 251, 27.7.2004, p. 9.

<sup>46</sup> "Implementing rules of Decision 2002/47/EC, ECSC, Euratom concerning the management of documents and of Decision 2004/563/EC, Euratom concerning electronic and digitised documents", SEC(2009)1643 of 30 November 2009.

<sup>47</sup> "Common Retention List for European Commission files", SEC(2007)970 of 4 July 2007 and "Common Retention List for European Commission files – first revision", SEC(2012)713 of 17 December 2012.

<sup>48</sup> SEC(2009)1643, section IV.1.

<sup>49</sup> The metadata of a document are: the date of the document; the date of registration; the registration number (unique identifier); the author of the document and, in the case of an individual, the department or body to which he/she belongs; the addressee of the document and, in the case of an individual, the department or body to which he/she belongs; where applicable, the level of security classification of the document and the security designators and markings; the title or subject of the document; the type of document concerned; the relevant heading of the filing plan and more specifically the file title; wherever possible, the number and a brief description of the attachments.

<sup>50</sup> Each heading in the filing plan must include the following metadata: title; brief description of scope; code; date of creation; date of closure; status; owner department and, where applicable, associated departments.

<sup>51</sup> Each file must include the following metadata: code of the heading in the filing plan; code of the file; title of the file; *chef de file* (lead department) for the file including, where applicable, successive *chefs de file* (lead departments); where applicable, level of security classification of the file; date of creation; date of closure; status of the file; content of a standard file, where applicable; electronic document management system, where applicable.

Among the sources of inspiration for this list (which aims to meet the needs of the institution) and for its implementation are ISO 15489 and MoReq, MoReq2 and MoReq2010, although it was necessary to interpret these sources in the light of the specific experience of document management at the Commission and to take this into account.

This list applies irrespective of the medium, digital or not, but in practice the role of metadata is considerably extended in the digital context. Of course, they serve first and foremost to describe and identify objects (such as documents, files etc.) in a manner understandable to users. But in an electronic document management system (EDMS), they are also directly used by the system itself, allowing for operations to be automated and rules implemented. The electronic system may also interact with metadata or develop new ones from the ones already existing.

In addition to a normative framework, the Commission has also developed a family of computer systems, called Hermes-Ares-Nomcom (HAN)<sup>55</sup>, for its document management. This has demanded a more detailed analysis of the metadata required by the legislative framework and their use, in order to get the most out of them, specifically in the light of their extended purpose and functioning in an EDMS.

---

<sup>52</sup> The preservation metadata for the files are the following: type of file; content, where appropriate; period of time the file has to be retained for administrative reasons (ARP); action to take when a file no longer holds administrative value (at the end of the ARP); action to be taken after transfer to the Historical Archives; *chef de file* (lead department) for preservation; location; medium; declassification, where appropriate; list of documents contained in the file.

<sup>53</sup> Following first review, in addition to the acquired metadata, the following metadata must be added to all transferred archives: date of transfer; identification of the transferring service; level of security classification; exceptions referred to under article 2 of Council Regulation No 354/83 as amended by Council Regulation No 1700/2003; expiration dates for the security classification or exceptions under article 2 of Council Regulation No 354/83 as amended by Council Regulation No 1700/2003 (where applicable).

<sup>54</sup> Following first or second review, the following acquired metadata must be preserved permanently in the references of eliminated files: all file codes (official, DG specific or other); title of the file; owner/lead department(s); date of creation; date of closure; file type; description of the file, including the number and type of archival units for paper files, or the comment 'electronic file' for electronic files, or the comment 'hybrid file' for files composed of both paper and electronic documents; to which the following metadata must be added: elimination date; department responsible for elimination.

<sup>55</sup> An overview of HAN can be found in White Paper 3 "ISO 30300-30301 – Système de gestion des documents d'activité : définition, modélisations, intégration aux autres normes de système de management", AFNOR, May 2012, pp. 34 f., accessible at <http://www.bivi.fonctions-documentaires.afnor.org/content/download/32987/197785/version/10/file/afnor-cn11+bilingue.pdf>.

It is not possible to list here all aspects of this work, nor to enter into much detail, but as an example, let us consider the metadata and functions related to the concept of the “closure date” of a file.

According to Commission rules the closure date is a mandatory metadatum of each file, defined by the IR as “the date when the most recent document was added to the file”<sup>56</sup>. However, this closure date is not necessarily the date on which the operation in the system of closing the file is carried out. The latter date is in fact an additional metadatum which is not required by the rules, but is automatically added by the system when a user closes the file (and is called “closing date” in order to avoid possible confusion between the two metadata).

To find the closure date, the system must then take into account the metadata of all documents that are filed in the file. These documents may be of two kinds: either they are just saved or they are registered.

In the current context a saved document is a document which does not fulfil the criteria for being formally and mandatorily registered, but which it has been decided to keep in the file for the sake of completeness. For a saved document, only the metadatum of title/subject must be entered by the user while the system also adds a save number and the date of creation/saving. A registered document must however be accompanied by all compulsory metadata on the above-mentioned list, including several dates: the date of creation/saving, the date of the document and the date of registration.

When determining the closure date for a file, the system must then compare all these dates and choose the most recent, which is then copied into the field “closure date”. From that date, and interaction with other sets of metadata, the system subsequently calculates a wide range of retention metadata for the file.

In fact, the above-mentioned CRL establishes an administrative retention period (ARP) and the post-ARP actions to be taken for each category of files. These data have been encoded in the system in the form of a list of CRL categories, enabling both a direct consultation of the CRL in the EDMS, and the (semi-) automatic use by the system of the CRL information. To each heading of the filing plan the appropriate categories of the CRL have been added, one of which is selected for each file under the heading (if only one category is available, it is automatically selected).

Once the file closure date has been calculated as described above, the system can proceed directly to the calculation and addition of *inter alia* the following file metadata: the date of the end of the ARP<sup>57</sup>, the post-ARP action to be undertaken (transfer to the historical archives, possibly after sampling or selection, or elimination) and the post-transfer action (permanent preservation or sampling/selection by the Historical Archives Service) in the event of a transfer.

On the basis of these metadata, manual or automated operations become possible, such as generating reports or alerting relevant users that files are ripe for transfer or elimination.

This example is only one amongst many of the use of metadata in the European Commission EDMS, but aims to illustrate the logic of how the system operates, while

---

<sup>56</sup> SEC(2009)1643, point II.4.6.1.

<sup>57</sup> This date is calculated from the closure date, by adding the number of years of the ARP and rounding up to the next whole year. For a file with the closure date of 1 May 2013 and a 5-year ARP, the end of the ARP is equal to 2019, and it is from the beginning of this latter year that the post-ARP action can be carried out.

showing that in the digital world, metadata provide numerous opportunities while at the same time they require rigorous attention and analysis. At the European Commission, this analysis was made both in advance (when drafting the IR) and incrementally during the operational development (as part preparing the specifications for and the development and implementation of the electronic tools). Throughout these analyses, standards (both formal, e.g. ISO 15489, and de facto, e.g. the three iterations of MoReq) often served as a source of inspiration and reference. However, in our experience we have never been able to copy or use these standards of the shelf. They have repeatedly been very useful to us, but adapted to our specific situation and to our practical needs.

### 13.1. THE IMPLEMENTATION OF ISO 23081 IN THE GOVERNMENT OF CANADA (GC) CONTEXT (NATHALIE VILLENEUVE)

---

In the GC, the Chief Information Officer Branch (CIOB) of Treasury Board of Canada Secretariat (TBS) is responsible for policy development governing GC institutions in the information management and information technology domains. TBS' information management policy framework includes the *Directive on Recordkeeping*, under which the *Standard on Metadata* is situated.<sup>[1]</sup> The former focuses on the management of information resources of business value (IRBV) within the broader context of information management, and the latter provides focused direction on how the management of IRBV is undertaken via metadata. Information resources (IRs) are defined as “any documentary material produced in published and unpublished form regardless of communications source, information format, production mode or recording medium.”<sup>[2]</sup> GC records are one type of IR, and IR's have business value when they are “created or acquired because they enable and document decision-making in support of programs, services and ongoing operations, and support departmental reporting, performance and accountability requirements.”<sup>[3]</sup> IRs with enduring (archival) value are a subset of IRBV.

The *Standard on Metadata* specifies requirements for three domains: recordkeeping, web resource discovery, and web content management systems. The recordkeeping metadata requirements are specified in Appendix B, and these are based on the generic set of metadata elements in *ISO 23081-2, Information and documentation – Records management processes – Metadata for records – Part 2: Conceptual and implementation issues*. Specific elements derived from the ISO generic set are identified in this appendix (s.3, s.4, s.5), and they document some aspects of records entities and agent entities as well as the event history. These elements must be applied by GC institutions in recordkeeping repositories. Additional metadata may be used in accordance with operational requirements. While

---

<sup>[1]</sup><http://www.tbs-sct.gc.ca/im-gi/imp-pgi/imp-pgi-eng.asp>

<sup>[2]</sup><http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=18909&section=text>

<sup>[3]</sup>Ibid.

technically the requirements in Appendix B could be interpreted as applicable to all types of IRBVs (not just records), the records-centric focus is clearly defined in the *Standard's* subordinate suite of recordkeeping metadata tools, the *GC Recordkeeping Metadata Element Set* (RKMES), *Recordkeeping Metadata Application Profile* (RKMAP) and *Recordkeeping Metadata Implementation Guide* (RKMIG). The RKMIG declares IRs such as published information and datasets as out of scope.<sup>[4]</sup>

The requirements stated in Appendix B of the *Standard* represent a partial reflection or implementation of ISO 23081-2 generic metadata elements within the ISO standard's conceptual metadata model. Not all of the generic set of metadata elements in ISO 23081-2 is present and, further, only a small portion of the conceptual model is represented. For example, only part of the hierarchical levels within the records and agents entity classes is specifically represented and other entity classes are absent altogether. In addition, relationships between entities are marginally represented and event planning metadata is absent. Part of the impetus behind this decision was to minimize the burden of onerous requirements on implementers, but also because a phased approach to defining metadata requirements was deemed desirable, to allow opportunity to benefit from experience gained through initial implementations and adaptations of the *Standard*.

So, while the *GC Standard on Metadata* is aligned with ISO 23081, it lacks comprehensive specificity with regard to both metadata elements to be implemented as well as how they are to be configured in relation to the ISO conceptual model. This limits the standardization of recordkeeping metadata per ISO 23081 in GC repositories. Gaps in coverage of TBS' *Standard on Metadata* in relation to recordkeeping requirements as per the *Directive on Recordkeeping* have been identified, and they will be addressed in subsequent revisions. For example, metadata elements specific to retention and disposition should be incorporated. There is also scope to clarify or broaden how ISO 23081 is to be reflected in GC institutions' recordkeeping metadata infrastructure in future versions of the *Standard*.

---

<sup>[4]</sup>Final drafts available at [http://www.gcpedia.gc.ca/wiki/Metadata\\_Tools](http://www.gcpedia.gc.ca/wiki/Metadata_Tools) (internal GC site). Publication pending, 2013.

## 13.2. VIEWPOINT ON CLOUD STORAGE (JOHANNE. LESSARD – QUEBEC)

---

In this context of shared and very often outsourced resources, cloud storage<sup>58</sup> acquires special importance when it comes to storing or preserving records. To fully understand the scope of cloud computing, it is necessary, first of all, to understand its practical foundations.

The servers belongs to the cloud services company.

- The application licences are paid for by the cloud services company.
- IT security is provided by the cloud services company.
- The data reside on servers belonging to the cloud services company.
- The servers are located on the premises of the cloud services company.
- Maintenance of the systems and applications is carried out by the cloud services company.
- The computers belong either to the organisation or the cloud services company.
- The data belong to the organisation.

Cloud storage has an advantage to which managers are particularly sensitive: an economic advantage. This is because not having to buy servers or product licences, not having to carry out IT maintenance and security and/or purchase and manage computer equipment generates substantial savings for an organisation. Many managers do not think twice about making the change.

However, cloud storage comes with its own set of risks.

- At a technical level, an organisation is no longer in control of the changes of version of the applications that it uses, which can cause a lack of integrity in documents.
- At a physical security level, an organisation no longer controls access to the premises in which its data is stored, which could give rise to unauthorised access to the organisation's data.
- In terms of accessibility to a company's data, an organisation generally has no right to inspect the employees of cloud services. This could give rise to a breach of confidentiality regarding sensitive data.
- In terms of the security of data transmissions, an organisation is at the mercy of the cloud service in the event of an outage. Every hour of production lost has an impact on the organisation's products/services.
- In terms of the physical location of the cloud service, the legal environment of the country in which this is located may be different to that of the organisation. This

---

<sup>58</sup>"Informatique nuagique" and "Infonuagique" are the French translations of cloud storage used in Canada.

may give rise to practices contrary to those of the legal environment in which the organisation operates and, in the event of a dispute (eDiscovery), the legal situation is unclear. Even though the data might belong to an organisation located in Canada for example (e.g.: not subject to the Patriot Act), if the data are stored in a cloud service in the US (subject to the Patriot Act), it would easily be possible to argue that data located on US soil is subject to US law, which does nevertheless not apply in the country in which the data originated. Because the legal apparatus is always behind technological advances, there is not yet a legal framework on this subject at an international level.

Before switching management of its data to a cloud service, an organisation's records and legal services manager should examine the following items:

- Protection and security of data.
- Continuity of business services.
- Security of data (integrity and reliability)
- Methods for retrieving data following termination of a contract.
- Issues concerning the transmission of data between countries/states/regions /districts.
- Transition methods (organisation/service or service/organisation)
- Service level agreement (availability and retrieval in the event of an incident)
- Confidentiality
- Price scales for each of the services provided.
- Intellectual property.
- Terms regarding consent to access an organisation's data.
- Audit methods.
- Taxes applicable to the transmission of data.
- Insurance.
- Liability and damages in the event of non-compliance with the contract drawn up.
- Governance (organisational and in the event of a dispute).

After analysis of all these risks, a positive business case for outsourcing to a cloud service is often more difficult to establish given an organisation's legal situation and area of business.

Cloud services have therefore developed the private cloud notion in order to diminish the adverse factors inherent to the risks. This alternative only partially diminishes all of the adverse factors, i.e. IT security, intellectual property and confidentiality; however, control metadata also need to be in place in order to ensure the security, integrity, reliability, confidentiality and "auditability" of data. It is therefore recommended that the management system for records used in a cloud environment be certified under the standards DoD 5015.2 or MoReq2, as these standards require audit metadata about the various aforementioned controls.

## 14 CONCLUSION

---

The revisions of the standards ISO 23081 and ISO 15489 in connection with the ISO 30300 standards on records management systems marks an important step in the development of the body of records management standards.

Firstly, these revisions are an opportunity to provide a reminder of and stress the need for metadata to be available as sets of information both to describe and manage records and their evolution.

Subsequently, although evidence has been provided of the usefulness of metadata, the fact remains that existing standards on locating, identifying and managing information and documents, in general, in systems, do not adequately satisfy records management requirements.

As a consequence, a principal - or a method - for preparing and developing metadata for records management needs to be devised in order to generate an operating model that can be adapted to the various situations that may be encountered in organisations. To this end and based on existing ISO standards and in particular ISO 15836, also known as Dublin Core, it does not appear that the reflection and formalisation work is complicated to carry out and complete. It should quickly result in the definition of a model built around a nucleus of metadata and specific layers given expansion rules that can cater to the diversity of records and the diversity of the existing management methods in organisations.

However, the question that now needs to be asked is how to position and integrate these developments into the body of standards and the technical reports connected with ISO 30300. Should there be wider adoption of the standard ISO 15489, which is very well understood and widely implemented by professionals, or should the standard ISO 23081 on metadata be revised? The economic pressure caused by the market due to cloud storage and big data only exacerbates the need and speeds up the obsolescence of existing standards reference systems.

In this context, we hope that this White Paper will contribute to the work that will be carried out during the ISO meetings due to take place in Paris during June 2013.

---

---

## 15 LE GROUPE DE TRAVAIL – WORKING GROUP

---

---

Ce Livre blanc a été conçu et rédigé à la demande des comités miroirs du TC46/SC11 de la Commission Nationale de normalisation 11 de l'AFNOR, de l'ILNAS Luxembourg et du Conseil Canadien des Normes :

*This White Paper has been designed and developed at the request of mirrors TC46/SC11 committees of the National Standards Commission 11 AFNOR, the ILNAS Luxembourg and the Standards Council of Canada:*

Ont contribué à la rédaction de ce Livre blanc / *Contributed to the drafting of the White Paper:*

- **Jacqueline Boga**, Information Management Officer, Faculty of Law, Queensland University of Technology. Brisbane, Australia.
- **Lucas Colet**, Ingénieur R&D, Centre de Recherche Public Henri Tudor, Président du Comité miroir Luxembourgeois ISO/TC46/SC11. R&D Engineer, Public Research Centre Henri Tudor, Chairman of Luxembourg Mirror Committee ISO/TC46/SC11.
- **Michel Cottin**, Coordinateur du groupe de travail Livre blanc, Groupe France Telecom –Orange, Records Manager Orange Labs, membre de l'AAF, Président du Comité miroir français ISO/TC46/SC11, membre de la Commission RM de l'AAF, administrateur de l'ADBS.
- **Cyrielle Demulier**, Master II Université d'Amiens, Stagiaire Records Manager Orange Labs.
- **Sylvie Forastier**, Central Filing & Archives Advisor Cabinet d'avocats international Linklaters, Luxembourg, Membre de la Commission Records Management de l'Association des Archivistes Français, membre du comité Luxembourgeois de ISO/TC 46/SC11.
- **Marine Fraiz**, Master II Université de Bourgogne. Stagiaire Records Manager, Orange Labs.
- **Mattias Hallin**, Archivist, European Commission, Secretariat-General, SG.DSG1.B.2 Document Management and Archives Policy - archiviste, Commission européenne, Secrétariat général, SG.DSG1.B.2 Politique documentaire et d'archivage
- **Lourdes Fuentes-Hashimoto**, Responsable des archives électroniques, ministère des Affaires étrangères, Direction des Archives
- **Johanne Lessard**, Analyste d'affaires et Experte de sujet chez Bombardier Aéronautique, Montréal, Québec, Canada et membre d'ARMA Montréal et International.

- **Charlotte Maday**, Responsable des archives, université Paris Diderot-Paris 7. Membre de la Commission Records Management de l'Association des Archivistes Français, membre du comité français de l'ISO/TC 46/SC11/CN11.
- **Philippe Martin**, Consultant indépendant spécialisé en gestion électronique de documents, Trésorier de l'APROGED, animateur de la commission normalisation, membre du TC 171 de l'ISO et de la CN 171 de l'AFNOR.
- **Katell Renaudie-Gueguen**, Consultante chez Locarchives, membre du Comité français de l'ISO TC/46/SC11, Animatrice du Groupe de Travail « procédures – modélisations – partage » de la Commission RM de l'AAF.
- **Paul Richy**, Expert en sécurité de l'information, Groupe France Telecom - Orange, Membre de l'ISO/IEC JTC 1/SC 27.
- **Marion Taillefer**, Chef du bureau de la bibliothèque du Conseil d'Etat, adhérente ADDBS, représentante de l'ADDBS auprès de la CG46/CN11 de l'AFNOR.
- **Katherine Timms**, Agent principal de projet, Direction d'évaluation de la pertinence du patrimoine documentaire, Bibliothèque et Archives Canada, membre de la délégation canadienne de l'ISO/TC46/SC11.
- **Nathalie Villeneuve**, Gestionnaire, Direction générale de l'évaluation de la pertinence du patrimoine documentaire, Bibliothèque et Archives Canada, membre du comité canadien de l'ISO/TC46/SC11.

**Gérard Dupoirier,**

Consultant du Cabinet « Document Numérique & Processus » a assuré la coordination éditoriale du Livre blanc.



**AFNOR**, Association française de normalisation.

En France, la normalisation, sur un plan général, bénéficie d'un statut fixé par décret dont la dernière version est le décret n° 2009-697 du 16 juin 2009 relatif à la normalisation. Il confirme les missions de service public de l'AFNOR, dont la représentation de la France dans des organismes

internationaux non gouvernementaux. Il confie, par ailleurs, au ministre chargé de l'Industrie, dont les attributions sont confiées au délégué interministériel aux normes, la définition de la politique des pouvoirs publics en matière de normes et le soin de s'assurer de la cohérence des actions des différents intervenants dans ce domaine. Il crée un groupe interministériel des normes (GIN) chargé d'assister le ministre dans la définition des orientations.



**AAF**, Association des archivistes français.

Elle défend les intérêts de la profession et cherche à promouvoir l'image et la visibilité du métier d'archiviste. Elle joue également un rôle actif dans les travaux de réflexion archivistique, en bonne intelligence avec le Service interministériel des Archives de France et les instances universitaires. L'AAF publie un bulletin bimestriel pour ses adhérents, *Archivistes !* et une revue scientifique, la *Gazette des archives*, destinée à tous

ceux qui s'intéressent, directement ou indirectement, aux archives et à la profession d'archiviste. Ses différents groupes de travail, dont la Commission Records Management Gestion des Documents d'Activité publient, par ailleurs et communiquent régulièrement leurs travaux sur le site internet, [www.archivistes.org](http://www.archivistes.org). Créé en 1984, le Centre de Formation de l'AAF a déjà accueilli chaque année près de 800 stagiaires. Animés par des professionnels du secteur, les stages, généralistes ou spécialisés, portent sur la gestion des documents d'activité, la théorie archivistique, la connaissance des institutions et les différentes techniques et actions mises en œuvre dans les services d'archives.



**ADBS**, Association des professionnels de l'information et de la documentation

Créée en 1963, l'ADBS, forte de ses 5 000 adhérents professionnels, est la première association professionnelle de l'information et de la documentation en Europe. Elle offre à ses adhérents les moyens de partager leurs expériences; de mutualiser leurs pratiques, d'échanger des services et actions et de constituer des réseaux. Elle a pour mission de promouvoir les métiers de la gestion de l'information et du document, de permettre aux professionnels de se maintenir au meilleur niveau

de compétences et de convaincre les pouvoirs publics de l'utilité de son action. Elle est présente tant au niveau national qu'en régions.



**APROGED**. Association de la Valorisation des Contenus numériques.

Créée en 1993, elle regroupe 120 adhérents représentant les professionnels de ce secteur en plein développement: dématérialisation, gestion de contenu et de documents, archivage électronique, capture, éditique, *workflow* indexation, linguistique, moteurs de recherche, sémantique, ... Tous les métiers sont présents: éditeurs, constructeurs, distributeurs, intégrateurs, prestataires de services, sociétés de conseil, tiers archiveurs, etc.

Au cœur de la réflexion sur la gestion et l'optimisation des contenus et des documents numériques, l'APROGED joue un rôle déterminant en matière d'information, d'étude, de formation, mais aussi de normalisation; l'association a pour objectif de faire des contenus (industriels, administratifs, culturels et scientifiques) l'atout majeur des entreprises et des organisations dans cette nouvelle gouvernance de l'information numérique en entreprise..



**ILNAS**, Institut Luxembourgeois de la Normalisation, de l'Accréditation, de la Sécurité et qualité des produits et services est une administration sous la tutelle du ministre ayant l'Economie dans ses attributions. Elle a été créée sur base de la loi du 20 mai 2008 et a démarré ses activités le 1er juin 2008. Pour des raisons de complémentarité, d'efficacité, de transparence et dans le cadre de la simplification administrative, l'ILNAS rassemble sous une même structure plusieurs missions administratives et techniques qui étaient auparavant dans les attributions de plusieurs structures publiques. Ces missions sont renforcées

et de nouvelles tâches sont également attribuées à l'ILNAS. L'ILNAS correspond à un réseau de compétences au service de la compétitivité et de la protection du consommateur. Une loi spécifique a été rédigée en ce sens et adoptée par la Chambre des Députés le 24/04/2008, constituant la feuille de route des missions à respecter par ILNAS.



## **Standards Council of Canada** **Conseil canadien des normes**

Le Comité canadien des normes (CCN) est le comité membre représentant le Canada à l'Organisation internationale de normalisation (ISO). La création

du comité consultatif canadien de la ISO/CG46 /CN11 a été approuvée par le CNN suivant l'intérêt des intervenants canadiens à participer dans ce champ d'activité ISO. Les comités consultatifs canadiens comprennent des individus compétents qui contribuent ensemble à une représentation équilibrée pour le développement de normes dans un domaine particulier. Pour plus d'information sur le rôle du Canada dans le développement de normes internationales ou autres activités de normalisation reliées, visitez le [www.scc.ca](http://www.scc.ca) or écrivez à [info@scc.ca](mailto:info@scc.ca).

The Standards Council of Canada (SCC) is the member body for Canada at the International Organization for Standardization (ISO). The formation of the Canadian Advisory Committee of ISO/TC 46/SC11 was approved by the SCC based on interest expressed by Canadian stakeholders in participating in this ISO field of activity. Canadian Advisory Committees are comprised of capable individuals whose collective input provides balanced representation on the development of standards in a specific subject area. For additional information about Canada's role in international standards development or other standardization-related activities, visit [www.scc.ca](http://www.scc.ca) or email [info@scc.ca](mailto:info@scc.ca).