View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by RERO DOC Digital Library

h e g
Haute école de gestion
Genève

La mesure des dimensions de la qualité des archives électroniques : apport des textes normatifs en matière d'archivage électronique à long terme

Travail de Bachelor réalisé en vue de l'obtention du Bachelor HES

par :

Julien TIÈCHE

Conseiller au travail de Bachelor :

Alain DUBOIS, Archiviste cantonal de l'État du Valais

Genève, 2 mars 2015

Haute École de Gestion de Genève (HEG-GE)

Filière Information Documentaire

Hes·so GENÈVE
Haute Ecole Spécialisée de Suisse occidentale

Déclaration

Ce travail de Bachelor est réalisé dans le cadre de l'examen final de la Haute école de gestion de Genève, en vue de l'obtention du titre de spécialiste en Information Documentaire.

L'étudiant a envoyé ce document par email à l'adresse remise par son conseiller au travail de Bachelor pour analyse par le logiciel de détection de plagiat URKUND, selon la procédure détaillée à l'URL suivante : http://www.urkund.fr/student_gorsahar.asp.

L'étudiant accepte, le cas échéant, la clause de confidentialité. L'utilisation des conclusions et recommandations formulées dans le travail de Bachelor, sans préjuger de leur valeur, n'engage ni la responsabilité de l'auteur, ni celle du conseiller au travail de Bachelor, du juré et de la HEG.

« J'atteste avoir réalisé seul le présent travail, sans avoir utilisé des sources autres que celles citées dans la bibliographie. »

Fait à Genève, le 2 mars 2015

Julien Tièche

Remerciements

Je tiens à remercier ici les différentes personnes qui m'ont apportées leur aide et leur support de quelque manière que ce soit. Merci tout spécialement à :

- Alain Dubois, mon conseiller, pour son aide dans la définition du sujet et les différentes réponses à mes questions ;
- Frédéric Noyer, des Archives de l'État de Neuchâtel, pour avoir accepté d'être mon juré ;
- Anouk Dunant Gonzenbach et Emmanuel Ducry, des Archives d'État de Genève, pour leur présentation de leur système d'archivage électronique à long terme ;
- Basma Makhlouf Shabou pour m'avoir orienté vers ce sujet ;
- Nicole Meystre Schaeren et l'ensemble de l'équipe des Archives de Montreux pour leur aide et leur flexibilité.

Résumé

Mesurer l'adéquation d'un produit ou d'un service, en regard d'exigences, d'attentes ou de besoins formulés par un individu ou imposés par un système institutionnel, est une activité qui doit être effectuée par la totalité des organisations. Le secteur des archives est également concerné par cette revue de la qualité. S'assurer de l'adéquation des documents d'activité et des archives est en effet la garantie, pour ces derniers, de répondre à des besoins probatoires, d'information, ou encore de recherche.

L'irruption des technologies numériques dans l'ensemble des activités humaines rend cette revue périodique de la qualité encore plus indispensable. En effet, dans un milieu en constante évolution, où l'obsolescence des systèmes de lecture des documents électroniques augmente le risque de perte ou de dégradation des archives, s'assurer du niveau adéquat de qualité permet de satisfaire aux exigences. Ainsi, les institutions en charge des archives ont la responsabilité de conserver des documents authentiques, fiables, intègres et exploitables.

Le but principal de ce travail est de livrer une réflexion théorique sur l'apport des normes et standards en matière d'archivage électronique à long terme dans la mesure des qualités des archives. A cette fin, trois textes normatifs ont été retenus : le modèle de référence OAIS, le standard d'encodage et de transfert de métadonnées METS, et enfin le dictionnaire de données pour les métadonnées de préservation PREMIS. Les éléments participant à la mesure des dimensions de la qualité des archives identifiés dans les normes et standards cités ci-dessus, ont été explicités et organisés selon un cadre conceptuel issu de la littérature existante sur le sujet. Afin de juger de leur pertinence, ils ont été associés à des variables et à des questions de mesure, puis opérationnalisés.

L'utilisation de normes et standards dans le milieu archivistique permet une compréhension entre professionnels, et est un prérequis à l'élaboration collective d'instruments et de méthodologies. Dans ce sens, les normes et standards sont un vecteur d'échange et d'automatisation. Identifier et expliciter des éléments participant à la qualité des archives dans un tel milieu contribuent ainsi grandement à l'automatisation de la mesure et donc à son adoption. L'accent de ce travail a par conséquent été mis sur l'identification d'éléments pouvant permettre une automatisation de la mesure, et par là même offrir un système applicable à de nombreuses organisations. Enfin, il propose des perspectives futures en termes de mesure, englobant d'autres normes archivistiques.

Table des matières

Déclaration i
Remerciements ii
Résumé iii
Liste des figures vi
1. Introduction 1
2. Méthodologie 4
3. État de l'art et définitions 6
3.1 Qualité 6
3.2 Dimensions de la qualité des archives 9
3.2.1 ISO 15489 Records Management 9
3.2.2 ISO 30300, 30301 Systèmes de gestion des documents d'activité 11
3.2.3 Étude sur la définition et la mesure des qualités des archives définitives issues d'une évaluation 12
3.2.4 Le projet QADEPs 17
3.2.5 Conclusion sur les dimensions de la qualité des archives 19
3.3 Archivage électronique à long terme 20
3.3.1 Contexte et problématique 21
3.3.2 Éléments pour la pérennisation de l'information 23
3.3.2.1 Modèle de référence OAIS 23
3.3.2.2 Principes et stratégies de pérennisation 25
3.3.3 Records Continuum et archives électroniques 27
3.3.4 Conclusion sur l'archivage électronique à long terme 29
4. Analyse des normes et standards 30
4.1 Aperçu sélectif des normes archivistiques 30
4.1.1 Les normes de description archivistique 31
4.1.2 Les normes de préservation 33
4.1.3 Les normes sur les documents d'activité 36
4.2 Choix des normes et standards analysés 36
4.3 OAIS 38
4.3.1 Origine et développement 38
4.3.2 Définition et principes 38
4.4 METS 42
4.4.1 Origine et développement 42
4.4.2 Définition et principes 42
4.5 PREMIS 43
4.5.1 Origine et développement 43
4.5.2 Définition et principes 43
4.6 Qualités identifiées 45

4.6.1 Dimension de la preuve crédible 46
4.6.1.1 Sous-dimension de l'authenticité..... 46
4.6.1.2 Sous-dimension de la fiabilité..... 49
4.6.1.3 Sous-dimension de la preuve historique 52
4.6.2 Dimensions de l'exploitabilité 53
4.6.2.1 Sous-dimension de l'accessibilité technique 54
4.6.2.2 Sous-dimension de l'accessibilité cognitive 57
4.6.2.3 Sous-dimension de l'accessibilité juridique 58
4.7 Conclusion sur les qualités identifiées dans les normes et standards...59
5. Opérationnalisation 61
5.1 Conclusion sur l'opérationnalisation.....73
6. Conclusion 75
Bibliographie 78
Annexe 1 : cadre conceptuel *étude sur la définition et la mesure des qualités des archives définitives issues d'une évaluation* 83
Annexe 2 : cadre conceptuel de la preuve crédible 84
Annexe 3 : cadre conceptuel de l'exploitabilité 85

Liste des figures

Roue de Deming 8
La dimension de la preuve crédible selon le projet QADEPs 18
La dimension de l'exploitabilité selon le projet QADEPs 19
La dimension de la représentativité selon le projet QADEPs 19
Flux de données et paquets d'information 24
Le modèle du Records Continuum 28
Étape du cycle de vie : *CREATE – PRESERVE - DESTROY*.....31
Le Matterhorn METS Profile et les standards de métadonnées archivistiques.....35
Les entités fonctionnelles du modèle OAIS 39
Les entités du modèle PREMIS.....45

1. Introduction

D'après l'Organisation internationale de normalisation, la qualité peut être définie comme "[l]'aptitude d'un ensemble de caractéristiques intrinsèques à satisfaire des exigences" (Organisation internationale de normalisation 2005, p. 7, 13). Dans un contexte archivistique, cette définition peut être comprise comme l'aptitude des archives à constituer un patrimoine documentaire représentatif d'une organisation et de son environnement, mais aussi à répondre à des besoins probatoires et d'information dans le cadre de la conduite d'activités ou d'obligations légales.

Comme pour l'intégralité des produits et des services, les archives et les documents d'activité[1] doivent tendre vers une amélioration permanente en termes de qualité. A cette fin, la qualité, ou les qualités des archives, sur la base de métriques préalablement définies, doivent pouvoir être mesurées (AFNOR 2012, p. 10). L'ensemble des processus archivistiques, de la création à la conservation, représente des investissements relativement importants. Dans ce sens, mesurer la concordance des archives, en regard de besoins ou d'attentes, permet de garantir la capacité de celles-là à remplir leurs fonctions tout au long de leur cycle de vie, et à justifier des investissements consentis dans les processus archivistiques. Ainsi, les choix en termes de conservation peuvent être légitimés, et les archives remplir leurs rôles spécifiques. La mise en place d'un tel système de management de la qualité permet de ce fait des actions correctives et est donc une composante de l'amélioration permanente en termes de qualité.

Dans un environnement électronique, où la masse documentaire est en pleine expansion et le décalage chronologique entre production et archivage non négligeable (Faivre d'Arcier 2004, p. 51), pouvoir s'assurer du niveau adéquat de qualité est une garantie supplémentaire en matière de conservation. En effet, la vulnérabilité des documents électroniques, du fait de l'obsolescence technologique et de la chaîne complexe de dispositifs de lecture requis (Banat-Berger, Duplouy, Huc 2009, p. 14, 15), rend obligatoire pour les archives électroniques de posséder certaines qualités concourant à leur pérennisation. Au-delà de cet aspect de permanence d'accès et de durabilité, l'intégrité, l'authenticité et la fiabilité des archives doivent aussi être garanties

[1] La norme ISO 30300 définit le document d'activité de la façon suivante : "informations créées, reçues et préservées comme preuve et actif par une personne physique ou morale dans l'exercice de ses obligations légales ou la conduite de son activité". Les archives sont définies, elles, comme "documents d'activité, conservés en vue d'une utilisation pérenne" (Organisation internationale de normalisation 2011a, p. 8, 9). Le terme document d'archives présent dans la norme ISO 15489 est un quasi synonyme de document d'activité (Organisation internationale de normalisation 2001a, p. 3).

(Kalbermatten 2012, p 33). C'est notamment au travers des métadonnées que ces différentes qualités peuvent être identifiées et mesurées. Les métadonnées ont une importance plus élevée dans un environnement électronique que dans un environnement papier, "car très peu de choses sont sous-entendues dans les systèmes électroniques et toutes les métadonnées implicites des documents papier doivent être explicitées" (Organisation internationale de normalisation 2001b, p. 8). Les systèmes normatifs sont d'une grande utilité en ce qui concerne les métadonnées.

L'omniprésence des normes et standards[2] dans le milieu des archives est telle que l'ensemble des fonctions archivistiques sont concernées. Que ce soit au niveau de la création, de la conservation ou encore de la diffusion, de nombreux textes normatifs viennent soutenir le professionnel dans son activité. Dans un environnement normatif, l'identification des éléments liés à la qualité des archives permet une mesure objective de ces dernières, et facilite les corrections éventuelles en amont de la chaîne de création des documents d'activité. De plus, l'identification de ces éléments est un préalable à l'automatisation de la mesure. En effet, la mise en exergue d'éléments normalisés et repérables sans ambiguïté, utilisés dans un grand nombre d'organisations, permet une mise en œuvre aisée de tests, de procédures de contrôles, et d'outils de mesure.

Ce travail a en premier lieu pour objectif de mener une réflexion théorique sur les liens entre, d'une part, les normes et standards en matière d'archivage électronique à long terme, et, d'autre part, les qualités des archives électroniques et leur mesure. A cette fin, une analyse de différents textes normatifs sera menée. Soit le modèle de référence OAIS, pour *Open Archival Information System*, connu aussi sous la norme ISO 14721, le standard d'encodage et de transfert de métadonnées *Metadata Encoding and Transmission Standard* (METS), ainsi que le dictionnaire de données pour les métadonnées de préservation *Preservation Metadata Implementation Strategies* (PREMIS). Les éléments qualitatifs ainsi identifiés seront organisés et explicités selon des modèles conceptuels issus de la littérature existante dans le domaine de la qualité des archives. L'accent sera principalement mis sur les éléments concourant à la mesure de l'intégrité, de l'authenticité, de la fiabilité et de l'accessibilité des archives.

[2] Une norme peut être définie comme un "document établi par consensus et approuvé par un organisme reconnu, qui fournit […] des règles, des lignes directrices […] garantissant un niveau d'ordre optimal dans un contexte donné" (Banat-Berger, Duplouy, Huc 2009, p. 35). De leur côté les standards sont élaborés par des communautés privées (Ibid.). Bien que le terme anglais "standard" désigne les deux types normatifs, les appellations de ce travail essaieront de préciser au maximum cette distinction.

A la suite de cette réflexion théorique, un volet opérationnel proposera des grilles de mesures des qualités identifiées. Dans cette dernière partie, la validité de ces grilles sera testée en regard d'éléments identifiables afin de juger de la reproductibilité de l'opérationnalisation, et donc de son automatisation.

2. Méthodologie

En septembre 2012, a débuté sous la direction de la professeure de la Haute École de Gestion de Genève (HEG) Basma Makhlouf Shabou, un projet intitulé *Définition et mesure des qualités des archives et documents électroniques* (Makhlouf Shabou 2013, p. 87). Celui-ci, mené en partenariat avec les Archives fédérales suisses, l'Université de Berne, les Archives d'État de Genève, les Archives d'État du Valais et la société Docuteam, avait pour objectif de fournir "un outil, clair et facilement utilisable, d'évaluation de la qualité des archives électroniques publiques dans la perspective de leur conservation pérenne dans un système de type OAIS" (Ibid.). A l'issu de ce projet, des variables permettant la mesure concrète des qualités des archives ont été établies et éprouvées.

Ce projet est né d'une thèse intitulée *Étude sur la définition et la mesure des qualités des archives définitives issues d'une évaluation* (Makhlouf Shabou 2013, p. 91). Cette thèse, également écrite par la professeure Basma Makhlouf Shabou, a voulu pallier à l'absence d'études dans le domaine de la définition et de la mesure des qualités des archives – ou des dimensions de la qualité des archives vu leur aspect pluriel – mais aussi proposer des méthodes et des instruments afin de permettre une opérationnalisation de la mesure (Makhlouf Shabou 2010, p. 6, 7).

Dans une approche similaire aux deux textes cités plus haut, le présent travail vise à mettre en avant les perspectives qu'offre la normalisation en matière de mesure des dimensions de la qualité des archives électroniques. L'accent sera particulièrement mis sur des éléments de mesure automatisables. A cette fin, il sera construit selon trois volets successifs : un volet théorique, un volet analytique et un volet pratique.

Le volet théorique aura pour mission d'identifier, de définir et de préciser les notions fondamentales permettant d'appréhender la problématique. Différents concepts seront ainsi précisés : le concept de qualité dans son acceptation générale, le concept de qualité des archives, et finalement celui d'archivage électronique à long terme. Si le concept de qualité sera présenté de façon très générale, sous son aspect de mesure notamment, celui des qualités des archives sera abordé à travers deux types de sources : les deux contributions explicites citées plus haut et les normes de système de management ISO 15489 (Organisation internationale de normalisation 2001a et b), ISO 30300 et 30301 (Organisation internationale de normalisation 2011a et b). Alors que les premières amènent une organisation des qualités des archives ainsi qu'un

cadre conceptuel, les deuxièmes préconisent un certain nombre de qualités, ou caractéristiques, que les documents d'activité ou les systèmes de gestion des documents d'activité se doivent de posséder. Enfin, les principaux concepts et notions relatifs à l'archivage électronique à long terme seront aussi mis en exergue. Il est en effet indispensable de disposer d'une base sur laquelle se reposer et de définir la problématique générale du sujet pour aborder ensuite la partie analytique.

Le volet analytique s'occupera, lui, d'identifier, de définir et d'organiser les éléments qualitatifs présents dans les textes normatifs sélectionnés. Dans ce but, ces derniers seront analysés et les qualités identifiées, organisées et explicitées selon les modèles conceptuels vus dans les recherches sur la qualité des archives. Les normes et standards retenus pour cette analyse seront aussi présentés dans cette partie. Un panorama succinct de l'environnement normatif de l'archivistique sera en outre proposé afin de situer ces derniers.

Troisième et dernier volet, le volet pratique aura pour mission l'opérationnalisation des qualités identifiées au préalable. Ces qualités et leurs variables associées seront mises en relation avec des indicateurs et des questions de mesure. Afin de permettre la mesure effective, des niveaux de qualité seront reliés à ces dernières et aux indicateurs. Le tout sera présenté sous forme de grilles permettant d'évaluer ces niveaux. Un test de faisabilité sera finalement opéré pour juger de la pertinence de cette opérationnalisation. Ce dernier se fera à l'aide d'instances XML intégrées aux grilles, où les éléments de mesures seront identifiés.

3. État de l’art et définitions

Le but premier de ce travail étant d'identifier et d'expliciter les éléments dans les normes et standards en matière d'archivage électronique à long terme participant à la mesure et à la validation des qualités des archives électroniques, une définition et une mise en contexte des concepts clés semble être un préalable. De façon plus précise, il s'agit de définir la notion de qualité et sa mesure, de rappeler les qualités des archives mises en exergue de façon implicite et explicite par la littérature spécialisée et les normes dites de management, et enfin, de définir les concepts liés à l'archivage à long terme des documents électroniques.

3.1 Qualité

Une explication générale de la notion de qualité semble être une bonne introduction pour aborder le concept de qualité des archives. Elle servira de base pour la définition des dimensions de la qualité des archives. Selon la norme ISO 9001, la qualité est "[l'] aptitude d'un ensemble de caractéristiques intrinsèques à satisfaire des exigences" (Organisation internationale de normalisation 2005, p. 7). Le terme intrinsèque est ici à comprendre comme antonyme d'attribué, et non pas comme l'exact opposé d'extrinsèque qui, lui, est défini comme dépendant de circonstances, de faits accessoires (Organisation internationale de normalisation 2005, p. 7 ; Centre National de Ressources Textuelles et Lexicales 2012). En réalité, les éléments extrinsèques ne sont pas fondamentalement exclus de cette définition. Au contraire, l'usage d'un produit ou son contexte de création participent également à la satisfaction des exigences.

Dans le champ d'activité de l'administration publique, l'Office québécois de la langue française (2014) définit la qualité comme "[l'] ensemble des caractéristiques d'un bien ou d'un service qui lui confèrent l'aptitude à satisfaire de manière continue les besoins et les attentes des utilisateurs ou des usagers". De façon plus précise, il s'agit de la capacité de traits distinctifs associés à un produit ou à un service, à répondre à des besoins ou des attentes, habituellement implicites ou imposés. Ces traits distinctifs ou caractéristiques peuvent être qualitatifs ou quantitatifs, de type physique, sensoriel, comportemental, temporel, ergonomique, ou encore fonctionnel (Organisation internationale de normalisation 2005, p. 7-13).

Delvosalle (2002), dans un propos axé sur l'administration publique, présente la qualité comme devant répondre à la "satisfaction de besoins réels". La qualité, écrit-il doit être "une mesure de l'harmonie de la rencontre entre les besoins des bénéficiaires, d'une

part et les résultats de l'organisation, d'autre part". En outre, en parlant de "non-qualité"[3], il la définit non seulement comme toutes choses pouvant être néfastes, mais aussi comme présentant des propriétés inutiles par rapport aux besoins exprimés. Cette approche permet de comprendre que la qualité ne doit pas forcément se situer à un niveau de "supériorité relative" (Delvosalle 2002), mais plutôt au niveau spécifique de la satisfaction de l'utilisateur. Les propriétés intéressantes, rappelle l'auteur, si elles ne sont pas demandées sont inutiles, et de surcroit peuvent engendrer des coûts supplémentaires (Ibid.).

La notion de qualité est aussi fortement liée à son contexte et à la façon dont elle est appréhendée. La qualité associe ainsi toujours le sujet (le client ou l'usager) et l'objet (le produit ou le service). Elle est donc constamment relative à un contexte, une attente (Chaumette 2006, p. 41). De même, elle a la particularité d'être multi-facette. C'est-à-dire qu'elle peut être identifiée et appréhendée selon plusieurs angles différents : aptitude à l'usage, aptitude à l'emploi, conformité aux exigences, ou encore satisfaction du client (Froman, Gourdon 2003, p. 155).

Un principe essentiel du management de la qualité[4] est son contrôle méthodique, la mise en œuvre de mesures correctives afin d'améliorer continuellement ses performances tout en répondant aux besoins des utilisateurs (Organisation internationale de normalisation 2005, p. v). Ces besoins pouvant évoluer avec le temps, une revue périodique des exigences de qualité est à prévoir (Froman, Gourdon 2003, p. 155).

En parallèle, il y a lieu de procéder à l'examen de la conception du produit ou service en question et de déterminer sa conformité en regard d'exigences spécifiées ou générales (Froman, Gourdon 2003, p. 96). Le contrôle méthodique de la qualité, et dès lors sa correction, doit donc être orienté sur deux axes : l'évolution des besoins, et la conception du produit ou du service, et par là même les caractéristiques qui leurs sont liées.

A cette fin, le management de la qualité doit mettre en place un système d'amélioration continue. Ce dernier peut être représenté, selon le responsable qualité de l'École polytechnique fédérale de Lausanne Michel Jaccard, sous la forme d'une roue de

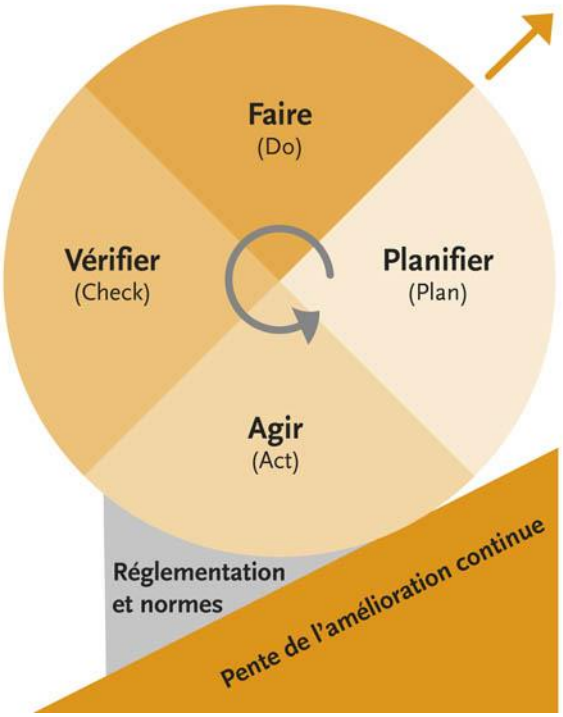
[3] La non qualité peut être comprise comme "tout échec, tout rebut, tout surcoût, tout retard, tout désagrément causé aux bénéficiaires, tout manquement quelconque à la mission de l'organisation" (Delvosalle 2002).

[4] Selon la norme ISO 9000 (Organisation internationale de normalisation 2008, p. 9), le management de la qualité est définit comme étant les "activités coordonnées permettant d'orienter et de contrôler un organisme […] en matière de qualité […]".

Deming (Jaccard 2010, p. 51). Ce dernier nous expose les quatre étapes de cette méthode itérative :

- "planifier (**P**lan) : formuler les objectifs et les processus nécessaires à l'obtention d'un produit conforme" (Ibid.) ;
- "faire (**D**o) : implanter le standard et les processus, ainsi que les mesures de contrôle" (Ibid.) ;
- "vérifier (**C**heck) : surveiller et évaluer les résultats obtenus ainsi que déterminer les actions correctives pour être en phase avec la conformité" (Ibid.) ;
- "corriger (**A**ct) : mettre en œuvre des actions correctives pour améliorer la performance des services et des produits" (Ibid.).

Roue de Deming



(http://www.enssib.fr/bibliotheque-numerique/notices/61230-schema-de-la-roue-de-deming)

En conclusion, la qualité est donc définie par la capacité d'un produit ou d'un service à satisfaire des besoins ou des attentes exprimés par des utilisateurs ou des usagers. Ces besoins et ces attentes peuvent être sous-entendus ou formulés et concerner des caractéristiques propres au produit ou au service, ou alors des caractéristiques extérieures à ces derniers. Cette qualité peut être appréhendée selon différents points

de vue et selon le contexte de l'objet et du sujet, que l'on se place au niveau du produit ou du service, de l'utilisateur et de ses attentes, ou encore d'exigences contractuelles ou règlementaires.

Du fait de la constante évolution des besoins des utilisateurs et de la nécessité de répondre à des exigences, le management de la qualité se doit de mettre en place des méthodes itératives et systématiques de contrôle. Ces dernières doivent participer à l'amélioration continue de la performance du management de la qualité, de la conformité du produit en regard d'attentes ou d'exigences, et donc des qualités qui lui sont liées.

3.2 Dimensions de la qualité des archives

La précédente section a permis d'appréhender le concept de qualité et sa mesure. Sur cette base, l'accent sera mis sur la qualité des archives de façon générale. Si, dans l'univers du management le concept de qualité peut se targuer de posséder une histoire relativement dense et lointaine[5], le concept de qualité dans le domaine de l'archivistique n'a que peu été étudié de façon formelle (Makhlouf Shabou 2010, p. 6). Qui plus est, il est assez récent. En effet, les deux seules contributions abordant de façon explicite et globale les qualités des archives remontent à la deuxième décennie du XXIème siècle. Si les apports explicites dans le domaine sont peu nombreux, la littérature normative identifie elle, de façon implicite, certaines caractéristiques ou éléments participant à la mesure des qualités des archives. Ainsi, les normes dites de système de management ISO 15489 sur le Records Management et ISO 30300, 30301 sur les Systèmes de gestion des documents d'activité, présentent les qualités des archives sous la forme de caractéristiques. Ces différents textes seront abordés dans la présente section, et les qualités qu'ils ont identifiées, répertoriées et explicitées.

3.2.1 ISO 15489 Records Management

La norme internationale ISO 15489 *Information et documentation – Records Management* "constitue un guide pour l'organisation et la gestion des documents d'archives" (Organisation internationale de normalisation 2001a, p. 1). Son champ d'application couvre donc les "documents créés, reçus et préservés à titre de preuve et d'information par une personne physique ou morale dans l'exercice de ses obligations légales ou la conduite de son activité" (Ibid., p. 3).

[5] En management, le concept de qualité découle du Taylorisme. Cette forme d'Organisation Scientifique du Travail est fixée dans les années 1880 par son inventeur Frederik Winslow Taylor (http://fr.wikipedia.org/wiki/Taylorisme [Consulté le 09.02.2015]). Le concept de qualité a été l'objet de nombreux écrits, démarches, et autres normes.

Comme indiqué plus haut, la norme internationale sur le Records Management ISO 15489 propose une approche implicite de la question de la qualité des archives. En effet, les qualités ne sont pas identifiées en tant que telles, mais plutôt sous la forme de caractéristiques. Ces dernières, au nombre de cinq, associées aux documents dont elles confèrent des traits distinctifs, doivent "permettre la continuité dans la conduite des affaires, satisfaire à l'environnement réglementaire et assumer leurs responsabilités" (Ibid., p. 6). Dans ce sens, les organismes se doivent de produire et conserver des documents présentant des caractéristiques d'authenticité, de fiabilité, d'exploitabilité, et d'intégrité. Et ce, aussi longtemps que nécessaire (Organisation internationale de normalisation 2001a, p. 6). A cette fin, la norme ISO 15489 propose cinq caractéristiques, ou qualités, que se doit d'avoir un document d'archives. Parmi ces dernières, quatre sont indépendantes de tous facteurs extérieurs et peuvent être qualifiées d'intrinsèques. La cinquième se rapporte à des éléments externes à l'objet et peut être étiquetée d'extrinsèque :

- **Authenticité** : "un document authentique est un document dont on peut prouver : a) qu'il est bien ce qu'il prétend être, b) qu'il a été effectivement produit ou reçu par la personne qui prétend l'avoir produit ou reçu, et c) qu'il a été produit ou reçu au moment où il prétend l'avoir été" (Organisation internationale de normalisation 2001a, p. 7). Pour garantir cette authenticité, les organismes doivent mettre en place des politiques et des procédures documentées, afin de contrôler les différentes interventions sur le document. Ces politiques et procédures doivent garantir l'identification des créateurs et la protection du document contre toutes interventions qui nuiraient à ce dernier. (Organisation internationale de normalisation 2001a, p. 7).
- **Fiabilité** : "un document fiable est un document dont le contenu peut être considéré comme la représentation complète et exacte des opérations, des activités ou des faits qu'il atteste, et sur lequel on peut s'appuyer lors d'opérations, d'activités ou de faits ultérieurs" (Organisation internationale de normalisation 2001a, p. 8).
- **Intégrité** : un document, pour être jugé intègre, doit être complet et non altéré. A cette fin, il doit être protégé contre toutes dégradations et autres falsifications (Organisation internationale de normalisation 2001a, p. 8).
- **Exploitabilité :** "un document utilisable est un document qui peut être localisé, récupéré, communiqué et interprété. Il convient qu'à chaque communication, le document soit relié à l'activité ou à l'opération à l'origine de sa création. Il convient que les liens contextuels des documents portent les informations nécessaires à la compréhension des opérations qui les ont crées [sic] et utilisées" (Organisation internationale de normalisation 2001a, p. 8).

A côté de ces quatre qualités, la norme ISO 15489 met en avant une cinquième caractéristique qui concourt grandement à leur renforcement : la traçabilité. Cette dernière qualité, centrée sur le système d'archivage, doit garantir la protection des documents contre toutes utilisations, transformations ou éliminations non autorisées

(Organisation internationale de normalisation 2001a, p. 10). La norme précitée la définit de la façon suivante :

- **Traçabilité** : la traçabilité est le fait de documenter tous les mouvements et toutes les utilisations d'un document (Organisation internationale de normalisation 2001a, p. 4).

Cinq qualités sont ainsi identifiées dans cette norme : l'**authenticité**, la **fiabilité**, l'**intégrité**, l'**exploitabilité** et la **traçabilité**.

3.2.2 ISO 30300, 30301 Systèmes de gestion des documents d'activité

La norme ISO 30300 *Information et documentation – Systèmes de gestion des documents d'activité* précise le vocabulaire utilisé dans la série des normes relatives aux systèmes de gestion des documents d'activité, de même que les objectifs et les bénéfices de ces systèmes. La norme ISO 30301 spécifie de son côté les exigences que doivent remplir un système de gestion des documents d'activité (Organisation internationale de normalisation 2011a, p. iv).

A l'instar de la norme ISO 15489, cette norme énonce de façon implicite certaines qualités que se doivent de posséder les documents d'archives. Un système de gestion des documents d'activité doit garantir "la création et le contrôle des documents d'activité répondant aux besoins de l'organisme et de la société, [et ce] aussi longtemps que nécessaire" (Organisation internationale de normalisation 2011a, p. 2). Afin de remplir ces objectifs, la présente norme demande la création et le contrôle de documents d'activité possédant différentes caractéristiques :

- **Fiabilité :** "un document d'activité est fiable si son contenu peut être considéré comme la représentation complète et exacte des opérations, activités ou faits qu'il atteste, et si l'on peut s'en servir lors d'opérations ou d'activités ultérieures. Il convient que les documents d'activité soient créés au moment de la transaction ou du fait auxquels ils se rapportent, ou juste après, par des personnes ayant une connaissance directe des faits ou par des outils utilisés systématiquement pour effectuer cette transaction" (Organisation internationale de normalisation 2011a, p. 3).
- **Authenticité :** pour qu'un document d'activité soit considéré comme authentique, il faut prouver qu'il correspond bien à ce qu'il devrait être, qu'il a été créé ou transmis par les personnes qui sont censées l'avoir créé ou transmis, et ce au moment supposé. Pour garantir l'authenticité des documents d'activité, il y a lieu de mettre en place des politiques et des procédures. Ces dernières doivent contrôler l'ensemble des interventions sur le document jusqu'à son sort final, et permettre l'identification des auteurs ainsi que la protection du document contre toutes interventions non autorisées (Organisation internationale de normalisation 2011a, p. 3).
- **Intégrité :** "l'intégrité d'un document renvoie au caractère complet et non modifié de son état" (Organisation internationale de normalisation 2011a, p.3). Comme indiqué ci-dessus, des politiques et procédures doivent être mises en place afin de protéger le document. Ces outils devront notamment préciser les

circonstances dans lesquels les documents d'activité peuvent être modifiés et qui sont les personnes habilitées à le faire. La norme précise encore sur ce point, que toute intervention sur un document d'activité doit être documentée et traçable (Organisation internationale de normalisation 2011a, p.3).

- **Utilisabilité :** "un document d'activité utilisable est un document qui peut être localisé, retrouvé, communiqué et interprété. Il doit pouvoir être présenté ultérieurement comme étant lié directement à l'activité ou à la transaction qui l'a produit. Les liens contextuels des documents d'activité doivent contenir les informations nécessaires à la compréhension des transactions qui les ont créés et utilisés. Il doit être possible de replacer les documents dans le contexte d'activités ou de fonctions élargies. Il convient que les liens entre les documents d'activité qui témoignent d'une succession logique d'activités soient maintenus" (Organisation internationale de normalisation 2011a, p.3).

Même si le domaine d'application de cette norme, tout comme la norme ISO 15489, est celui des documents d'activité, les qualités qu'elle souligne doivent être prises en compte et exploitées dans une optique de définition et de mesure de la qualité des archives de façon générale. Les archives, nous rappelle la norme ISO 30300, sont les documents d'activité "conservé[s] en vue d'une utilisation pérenne" (Organisation internationale de normalisation 2011a, p. 8). Elles sont conservées de façon définitive pour leur valeur d'information et de témoignage, au détriment d'autres documents d'activité, détruits, car ne possédant pas ces valeurs. Les caractéristiques que se doivent de posséder les documents d'activité, doivent par conséquence perdurer dans leur statut d'archives.

Qui plus est, les deux normes dites de système de management se placent dans une démarche qualité (AFNOR 2012, p. 22), et par là même participent grandement à notre propos. Quatre qualités se dégagent ainsi de cette norme : la **fiabilité**, l'**authenticité**, l'**intégrité** et l'**utilisabilité**. Cette dernière qualité peut être rapprochée de l'exploitabilité vue dans la norme précédente. Une cinquième qualité, la **traçabilité** apparaît de façon sous-jacente à l'**intégrité**.

3.2.3 Étude sur la définition et la mesure des qualités des archives définitives issues d'une évaluation

La première contribution explicite et globale à l'étude et à la définition des qualités des archives remonte à 2010 (Makhlouf Shabou 2010, p. 6). Avec sa thèse de doctorat intitulée *Étude sur la définition et la mesure des qualités des archives définitives issues d'une évaluation*, Basma Makhlouf Shabou a voulu pallier l'absence de recherches dans le domaine de la définition et de la mesure des qualités des archives, mais aussi proposer des méthodes et des instruments afin de permettre une opérationnalisation de la mesure de la qualité des archives (Ibid., p. i, ii).

Cette étude s'est déroulée en deux phases : une phase conceptuelle et une phase empirique. La première phase a permis d'identifier, de définir et d'organiser les dimensions suivantes de la qualité des archives : l'**unicité**, la **preuve crédible**, l'**exploitabilité** et la **représentativité**. Les deux premières sont intrinsèques, tandis que les deux suivantes sont extrinsèques (Makhlouf Shabou 2010, p. 112-126, 129). Toutes ont été mises en exergue grâce à une revue approfondie de la littérature spécialisée qui a porté sur quatre objets principaux : la fonction d'évaluation, les principes qui en découlent[6], les approches d'évaluation[7] et la qualité des archives en archivistique contemporaine (Ibid., p. 11). Basma Makhlouf Shabou a ensuite élaboré un cadre conceptuel permettant d'organiser et de définir les dimensions des qualités des archives. Ce dernier est composé de quatre niveaux fondés sur une structure hiérarchique allant du général au particulier : dimensions, indicateurs généraux de niveaux 1, indicateurs généraux de niveaux 2, et indicateurs spécifiques. (Ibid., p. 110 ; annexe 1).

Ainsi, les dimensions identifiées de la qualité des archives se sont vues attribuer des niveaux conceptuels hiérarchiques précisant leur nature. Une partie de ces dimensions et leurs niveaux respectifs sont résumés plus loin. Le présent travail se focalisant particulièrement sur les qualités identifiables et mesurables des archives dans un environnement électronique, ce résumé aura trait uniquement à deux dimensions, la preuve crédible et l'exploitabilité.

La **preuve crédible**, qui est une dimension intrinsèque de la qualité, est déterminée par la "capacité d'un document à gagner la confiance de son utilisateur comme étant la source privilégiée à l'appui des faits" (Makhlouf Shabou 2010, p. 115). Elle repose sur l'authenticité et la fiabilité des documents selon les normes ISO 15489 et ISO 30300. Un document d'activité authentique doit prouver qu'il correspond bien à ce qu'il devrait être, qu'il a été créé ou transmis par les personnes qui sont censées l'avoir créé ou transmis, et ce au moment supposé (Organisation internationale de normalisation 2011a, p. 3). Sous l'authenticité, indicateur général de niveau 1, l'auteure définit deux indicateurs de niveau 2, soit l'identité exclusive et l'intégrité du format (Makhlouf Shabou 2010, p. 116) :

Identité exclusive

[6] Respect des fonds, cycle de vie des documents, théorie des valeurs.

[7] Documentation Plan, Documentation Strategy, macro-évaluation, approche centrée sur l'utilisateur, approche de l'évaluation intégrée.

Ensemble de caractéristiques d'un document ou document d'activité qui l'identifie de façon unique et qui le distingue des autres (InterPARES 2 Project 2012).

Statut des pièces du dossier

"Le statut des pièces du dossier est lié à leur nature (original, copie authentifiée, copie, brouillon) et à leurs différents signes particuliers (confidentiel, traduction etc.) et d'authenticité" (InterPARES 2 Terminology Database 2010, cité dans Makhlouf Shabou 2010, p. 116).

Intégrité du format

État complet et non altéré (InterPARES 2 Project 2012).

Transmission

Mouvement d'un document d'activité à travers l'espace (entre personnes physiques, personnes morales, ou systèmes) ou le temps (InterPARES 2 Project 2012).

Préservation

"The whole of the principles, policies, rules and strategies that controls the physical and technological stabilization and protection of the intellectual form of acquired records intended for their continuing, enduring, stable, lasting, uninterrupted and unbroken chain of preservation, without a foreseeable end" (InterPARES 2 Terminology Database 2010, cité dans Makhlouf Shabou 2010, p. 117).

Conservation

Ensemble des outils régissant la tenue et la garde des archives, ainsi que tous les aspects pouvant être liés à leur protection et à leur état matériel (Makhlouf Shabou 2010, p. 117).

Deuxième indicateur de niveau 1 pour la dimension de la preuve crédible, la fiabilité est selon la norme ISO 15489, la capacité pour un document à être considéré du point de vue de son contenu, comme "la représentation complète et exacte des opérations, des activités ou des faits qu'il atteste, et sur [laquelle] on peut s'appuyer lors d'opérations, d'activités ou de faits ultérieurs" (Organisation internationale de normalisation 2001a, p. 8). Sous cet indicateur général de niveau 1, deux indicateurs de niveau 2 ont été définis, soit le contexte de production ainsi que la complétude (Makhlouf Shabou 2010, p. 117, 118) :

Contexte de production

"Éléments essentiels qui caractérisent la création des archives" (Makhlouf Shabou 2010, p. 118).

Créateur

Personne physique ou morale, qui créée, reçoit ou gère des documents d'activité du fait de ses mandats, missions, fonctions ou activités (InterPARES 2 Project 2012).

Procédure de création

Procédure à l'origine de la formation du document d'activité, ou participant à sa création (InterPARES 2 Project 2012).

Complétude

"The characteristic of a record that refers to the presence within it of all the elements required by the creator and the juridical system for it to be capable of generating consequences" (InterPARES 2 Terminology Database 2010, cité dans Makhlouf Shabou 2010, p. 118, 119).

Contenu

"Message contenu dans un dossier" (InterPARES 2 Terminology Database 2010, cité dans Makhlouf Shabou 2010, p. 119).

La dimension de l'**exploitabilité**, qui est de son côté une dimension extrinsèque, est selon l'auteure "tout aspect susceptible de faciliter et de promouvoir l'utilisation des archives définitives" (Makhlouf Shabou 2010, p. 120). Cette dimension de la qualité repose sur trois indicateurs généraux, soit l'accessibilité cognitive des archives, leur accessibilité juridique et leur accessibilité matérielle.

Accessibilité cognitive

Accès aisé au contenu et au contexte de création des archives (Boles et Young, 1991, cité dans Makhlouf Shabou 2010, p. 120).

Identifiabilité

"Possibilité et facilité d'identifier des archives à partir de deux éléments, soit d'une part la structure classificatoire logique qui sert à organiser intellectuellement les documents (Schellenberg, 1965, cité Makhlouf Shabou

2010, p. 121), tel le plan de classification, et d'autre part le titre attribué au dossier d'archives" (Makhlouf Shabou 2010, p. 121).

Compréhensibilité

Capacité et facilité à appréhender le contenu des archives. La compréhensibilité repose sur l'organicité et le format de représentation des archives (Makhlouf Shabou 2010, p. 121).

Contextualité

"Compréhension et accès aux différents éléments du contexte de création et d'utilisation des archives (Senécal, 1991 ; Menne-Haritz, 1994, cité dans Makhlouf Shabou 2010, p. 121) à partir des descriptions disponibles" (Makhlouf Shabou 2010, p. 121).

Accessibilité juridique

Possibilité de communiquer les archives sur la base de textes juridiques ou de règlements (Makhlouf Shabou 2010, p. 121).

Communicabilité légale

Possibilité de communiquer les archives garantie par des textes juridiques (Makhlouf Shabou 2010, p. 121).

Communicabilité réglementaire

Possibilité de communiquer les archives garantie par des textes réglementaires (Makhlouf Shabou 2010, p. 121).

Accessibilité matérielle

"Facilité d'accéder au support des archives définitives. Elle repose sur la repérabilité physique et la lisibilité du support" (Makhlouf Shabou 2010, p. 122).

Repérabilité

Capacité et facilité à localiser les archives dans un dépôt de conservation, et à retracer leur mouvements (Organisation internationale de normalisation 2001b, cité dans Makhlouf Shabou 2010, p. 122 ; Makhlouf Shabou 2010, p. 122).

Lisibilité

"Possibilité de lire les documents grâce à la qualité de leur état physique ainsi que la disponibilité du matériel approprié à leur lecture" (Makhlouf Shabou 2010, p. 122).

Dans la phase empirique de son étude, l'auteure a cherché à opérationnaliser, au travers de variables résultants des indicateurs spécifiques, les deux dimensions suivantes de la qualité des archives : la preuve crédible et l'exploitabilité. Cette opérationnalisation s'est faite en associant à ces variables des questions et des niveaux de mesure. Pour cela, elle a collecté auprès de l'institution Bibliothèque et Archives nationales du Québec des données comprenant notamment un panel de fonds représentatifs situés au centre d'Archives de Montréal (Makhlouf Shabou 2010, p. 110, 128). Cette collecte de données, faite sous la forme d'une grille de mesure, a permis de valider l'opérationnalisation de différents indicateurs des deux dimensions de la qualité citées plus haut, et de mettre en évidence les points forts et les faiblesses des lots examinés (Ibid., p. 213, 228).

L'apport de cette première contribution explicite dans le domaine de la qualité des archives est multiple. Premièrement, elle offre un cadre conceptuel permettant de définir le concept de qualité des archives définitives. Deuxièmement, elle propose une méthode et des instruments de mesure des qualités des archives. Enfin, elle permet de vérifier et de valider les décisions liées à l'une des principales fonctions archivistiques, l'évaluation, et ainsi de justifier les investissements, aussi bien matériels que financiers ou humains consentis (Makhlouf Shabou 2010, p. 236, 237). Ce dernier point revêt une importance capitale, car il permet de renforcer la crédibilité du métier d'archiviste.

Quatre dimensions de la qualité des archives ont été identifiées lors de ce travail : l'**unicité**, la **preuve crédible**, l'**exploitabilité**, et la **représentativité**. Sous ces dernières, différents indicateurs ont été définis. Ces derniers devant permettre l'explicitation conceptuelle de la dimension et permettre ainsi la mesure (Makhlouf Shabou 2010, p. 111).

3.2.4 Le projet QADEPs

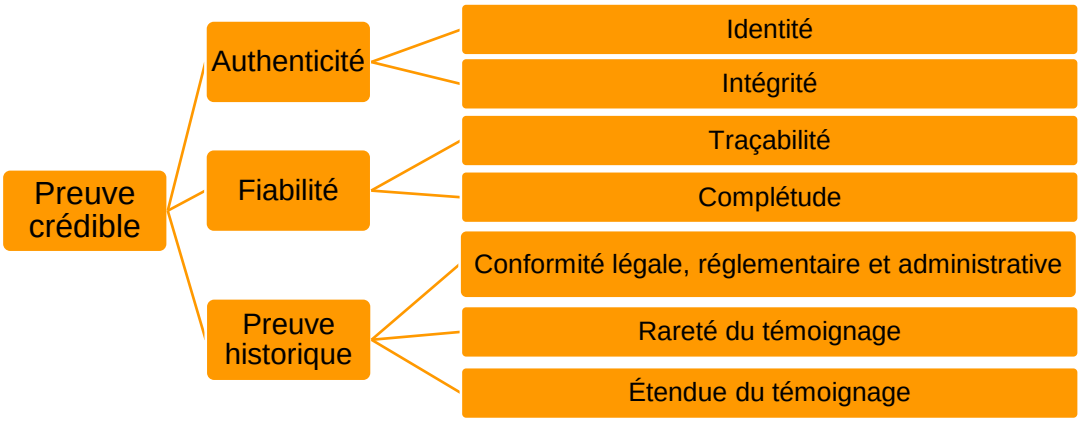
Sous l'acronyme QADEPs pour *Définition et mesures des qualités des archives et documents électroniques publics*, un second travail de la même auteure vue plus haut, tente d'apporter sa contribution sur le sujet de la qualité des archives. L'objectif de ce projet est de proposer aux professionnels un instrument, facile d'utilisation, d'évaluation de la qualité des archives électroniques publiques dans une optique de préservation durable dans un système de type OAIS (Makhlouf Shabou 2013, p. 87). Regroupant

plusieurs partenaires suisses publics ou privés[8], cette contribution s'appuie sur une organisation similaire au travail de recherche vue plus haut : définition d'un cadre conceptuel réunissant les dimensions et indicateurs de la qualité et développement de variables permettant la mesure de ces dernières. Sous ce dernier point, diverses cibles ont été définies : favoriser l'automatisme des mesures, restreindre le caractère subjectif de la mesure, favoriser la réutilisabilité de l'étude, et s'assurer du caractère applicable et pertinent des variables définies au travers de tests dans des institutions (Ibid., p. 88).

Trois dimensions, huit sous-dimensions et dix-sept indicateurs de la qualité des archives électroniques ont été mis en exergue au travers du cadre conceptuel (Makhlouf Shabou, p. 91). L'organisation de ces éléments est résumée ci-dessous :

Preuve crédible : "capacité d'un document à gagner la confiance de son utilisateur comme étant la source privilégiée à l'appui des faits" (Makhlouf Shabou 2010 ; InterPARES 2 Terminology Database 2010, cité dans Makhlouf Shabou 2013, p.91).

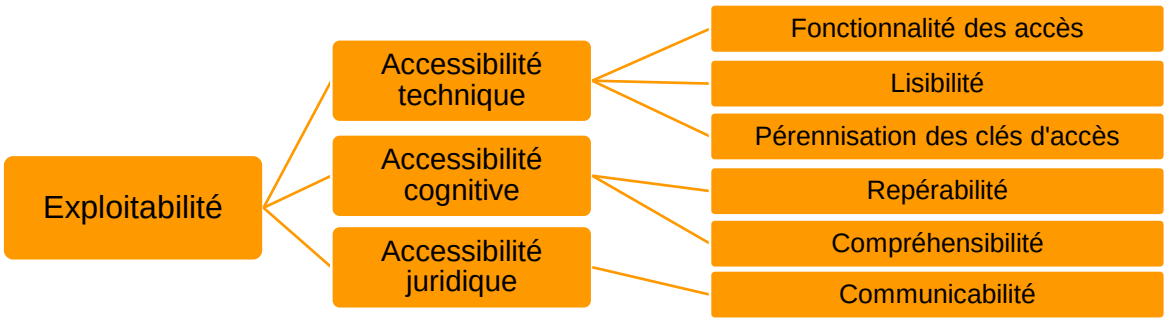
La dimension de la preuve crédible selon le projet QADEPs



[8] Archives fédérales suisses, Université de Berne, Archives d'État de Genève, Archives d'État du Valais, Docuteam GmbH.

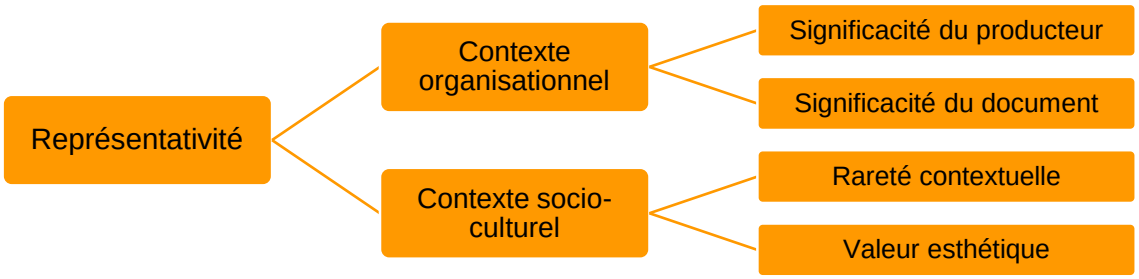
Exploitabilité : "un document jugé exploitable […] est un document utilisable qui peut être localisé, récupéré, communiqué et interprété aisément" (Organisation internationale de normalisation 2001, p. 8, cité dans Makhlouf Shabou 2013, p.92).

La dimension de l'exploitabilité selon le projet QADEPs



Représentativité : "capacité des archives à permettre un témoignage significatif, riche et exhaustif des différents éléments du contexte organisationnel de leur création" (Makhlouf Shabou 2010, p. 123, cité dans Makhlouf Shabou 2013, p.92).

La dimension de la représentativité selon le projet QADEPs



3.2.5 Conclusion sur les dimensions de la qualité des archives

Les différents textes vus plus haut permettent d'appréhender de façon globale le concept de qualité des archives. Les textes normatifs ISO 15489, ISO 30300 et ISO 30301 proposent des qualités que se doivent d'avoir les documents d'activité, et donc

les archives au sens large, alors que les deux contributions explicites apportent une vue conceptuelle et hiérarchique de ces différentes qualités, et permettent par là même la création d'indicateurs et de variables permettant la mise en œuvre de la mesure.

L'ensemble de ces contributions est une base solide pour l'identification, l'explicitation et l'opérationnalisation d'éléments concourant à la mesure des dimensions de la qualité des archives électroniques présents dans les normes et standards en matière d'archivage électronique à long terme.

3.3 Archivage électronique à long terme

Cette section a pour but de définir et de mettre en exergue certains concepts et notions relatifs à l'archivage électronique à long terme. Elle n'a pas pour objectif d'expliciter les normes et les standards choisis pour cette étude, ces derniers étant énoncés plus loin, mais de définir la problématique générale du sujet, ainsi que quelques principes et stratégies de pérennisation. Cependant, pour préciser les concepts de base, la norme OAIS y sera abordée de façon générale.

Avant d'aborder concrètement la problématique du sujet, il semble utile de revenir rapidement sur des questions de vocabulaire et de synonymie. En effet, si le nom choisi pour ce travail est celui d'archivage électronique à long terme, la revue de la littérature dans le domaine nous propose différentes appellations de même sens ou de sens voisin. Selon le Digital Preservation Coalition (DPC), le terme de *digital preservation* se réfère à l'ensemble des activités indispensables à assurer un accès continu à l'information numérique aussi longtemps que nécessaire. Toujours selon le même groupe, ce nom est utilisé – dans le domaine archivistique et bibliothéconomique – en lieu et place de *digital archiving* (Digital Preservation Coalition, 2012). "Long-term digital preservation are continuous process and technologies which guarantee that archived digital information remains accessible and usable for long-time periods that span over several decades and even centuries" (Brunsmann et al. 2012, p. 27). A travers ces définitions, nous disposons déjà de quelques éléments quant à la signification de l'archivage électronique, et de son implication sur le long terme. Le but de l'archivage électronique à long terme est donc de maintenir accessible et utilisable les informations numériques, et ce sur une durée supérieure à des décennies voire des siècles. Une conception de ce "long terme" peut être "défini[e] comme étant suffisamment long pour être soumis à l'impact des changements technologiques, y compris à la prise en compte de nouveaux supports et nouveaux formats de données ou à des changements de la communauté d'utilisateurs" (Comité consultatif pour les systèmes de données spatiales 2005, p. 1-1). Comme

énoncé plus haut, différents termes sont utilisés en anglais pour recouvrir les notions qui nous intéressent : il s'agit principalement de *digital preservation*, *digital archiving*, *long-term digital preservation*, voire *long-term digital archiving*. A noter que le terme *digital curation* est utilisé avec un sens proche dans le domaine de la recherche : "Digital curation involves maintaining, preserving and adding value to digital research data throughout its lifecycle" (Digital Curation Centre, 2004-2015). Dans notre langue, si le nom choisi pour ce travail est celui d'archivage électronique à long terme, le nom d'archivage numérique à long terme est fréquemment utilisé en lieu et place de celui-là. En allemand pour finir, le terme de *Digitale Langzeitarchivierung*[9] est le plus usité.

3.3.1 Contexte et problématique

Depuis quelques années, un formidable développement des technologies numériques se fait sentir sur l'ensemble de nos activités. "Cette évolution a pour conséquence que toute l'information, dans notre société, est créée sous cette forme, et quand ce n'est pas le cas, comme pour un document papier, on procédera dans un certain nombre de cas à une numérisation" (Banat-Berger, Duplouy, Huc 2009, p. 14). Deux principales raisons à cette irruption numérique dans notre quotidien : augmentation de la puissance, de la performance et de l'offre des ordinateurs, logiciels, réseaux, systèmes de stockage et autres services dématérialisés d'une part, et diminution des coûts d'autre part (Ibid., p. 10). Ce formidable développement à des effets en termes de durée de vie des technologies numériques. Du fait de la concurrence économique, les produits issus des industries remplacent et rendent obsolète leurs prédécesseurs. Cette obsolescence du matériel et des logiciels informatiques a des conséquences sur la conservation et l'exploitation des documents électroniques. En effet, lorsqu'un dispositif de lecture est frappé d'obsolescence, les données qu'il devrait interpréter deviennent inutilisables (Trace, Sanett 2000, p. 24). La durée de vie d'un document électronique serait ainsi d'une dizaine d'année s'il n'a fait l'objet d'aucune mesure particulière (Banat-Berger, Duplouy, Huc 2009, p. 15). A noter qu'au-delà de cette notion "d'obsolescence programmée"[10], une autre menace pèse sur l'information numérique : la dégradation des supports de stockage. Effectivement, les supports de stockage, quelle que soit leur forme, sont soumis à des menaces dont les causes peuvent être physiques, chimiques, environnementales ou organisationnelles (Banat-

[9] Le nombre pondéré de pages indexées sous les différents termes par le moteur de recherche Google est révélateur : 1'530 occurrences pour le terme "archivage électronique à long terme" et 2'440 pour le terme "archivage numérique à long terme". En allemand, 71'100 occurrences pour "digitale Langzeitarchivierung" et 2'160 pour "elektronische Langzeitarchivierung" [Consulté le 16.02.2015].

[10] L'obsolescence programmée est le fait de réduire la durée de vie d'un produit afin d'amener le consommateur à en acheter de façon prématurée un nouveau (http://obsolescence-programmee.fr/ [Consulté le 09.02.2015]).

Berger, Duplouy, Huc 2009, p. 84). Cette perte de données est d'autant plus probable que le dispositif de lecture est le plus souvent complexe et repose sur un ensemble d'éléments[11] dont "l'indisponibilité d'un seul […] au sein du dispositif de lecture peut compromettre définitivement un patrimoine informationnel" (Banat-Berger, Duplouy, Huc 2009, p. 15). Au-delà de garantir la longévité et l'accessibilité des données électroniques, il s'agit aussi de garantir l'intégrité et l'authenticité de ces dernières (Kalbermatten 2012, p. 33). En effet, les trois plus importantes charges d'un système d'archivage électronique sont de garantir l'authenticité, la fiabilité et l'intégrité des objets numériques, de vérifier que les objets conservés n'ont pas été endommagés, et enfin, de planifier la pérennisation du contenu en regard de l'obsolescence technologique (National Information Standards Organisation p. 62).

Conserver et maintenir l'ensemble des éléments du dispositif de lecture sous la forme d'un musée n'est pas une solution adéquate, du fait de l'accélération des innovations et des changements fréquents de technologies y résultant (Brunsmann et al. 2012, p. 27). Le transfert, à des fins de conservation, sur supports analogiques des documents numériques, n'est pas non plus une solution viable. Tout d'abord, les pertes de fonctionnalités résultant d'une telle opération ne sont pas acceptables, notamment en ce qui concerne les documents polymorphes[12]. Ensuite, certains types de documents, où la notion de page n'a aucun sens, ne peuvent être transférés sur papier. C'est le cas entre autres des documents multimédia. Pour finir, "les quantités colossales de données […], les nécessités impérieuses de pouvoir traiter ces données avec des techniques d'analyse numérique de plus en plus efficaces excluent tout recours aux supports analogiques" (Banat-Berger, Duplouy, Huc 2009, p. 14).

Cette problématique d'obsolescence des technologies numériques ne concerne pas que les documents conservés pour des raisons patrimoniales, mais bien l'ensemble des archives électroniques. En effet, certains secteurs d'activité se doivent de conserver les données relatives à un produit, pour des besoins légaux ou économiques, au moins aussi longtemps que ce dernier sera en fonctionnement (Brunsmann et al. 2012, p. 28). La difficulté d'une telle contrainte est claire lorsque l'on sait que les logiciels permettant l'interprétation de ces données ont une durée de vie inférieure aux obligations de conservation de ces données (Ibid., p. 27). Pour exemple, la durée de vie d'un avion peut être supérieure à septante ans. Les données y

[11] Ces éléments sont multiples : ordinateurs, systèmes d'exploitation, logiciels, formats de données…

[12] Les formulaires PDF sont un exemple de documents polymorphes (Banat-Berger, Duplouy, Huc 2009, p. 13).

afférentes devront donc être conservées pour une durée pouvant excéder 99 ans (Ibid. p. 30).

La problématique nous concernant ici en est encore accentuée du fait de l'organisation en couche de l'information. En effet, pour rendre intelligible des bits stockés sur un support physique, un ensemble de règles et d'algorithmes sont nécessaires. Ces derniers peuvent être représentés sous la forme de couches. Chaque couche "répondant à des fonctionnalités distinctes" (Banat-Berger, Duplouy, Huc 2009, p. 26-32). La connaissance de ce modèle de couches, ou format de données, est un gage essentiel pour la conservation. Cette connaissance de la spécification des formats de données va permettre entre autre d'évaluer les conséquences de la migration de formats et de développer des outils de restitution (Ibid., p. 32).

3.3.2 Éléments pour la pérennisation de l'information

Comme vu précédemment en début de ce chapitre, le but de l'archivage électronique à long terme est de maintenir accessible, utilisable, intègre, fiable et authentique les informations numériques, et ce sur une durée suffisamment longue pour subir les changements technologiques tels que l'arrivée de nouveaux supports, formats de données, ou encore des modifications de la communauté d'utilisateurs.

3.3.2.1 Modèle de référence OAIS

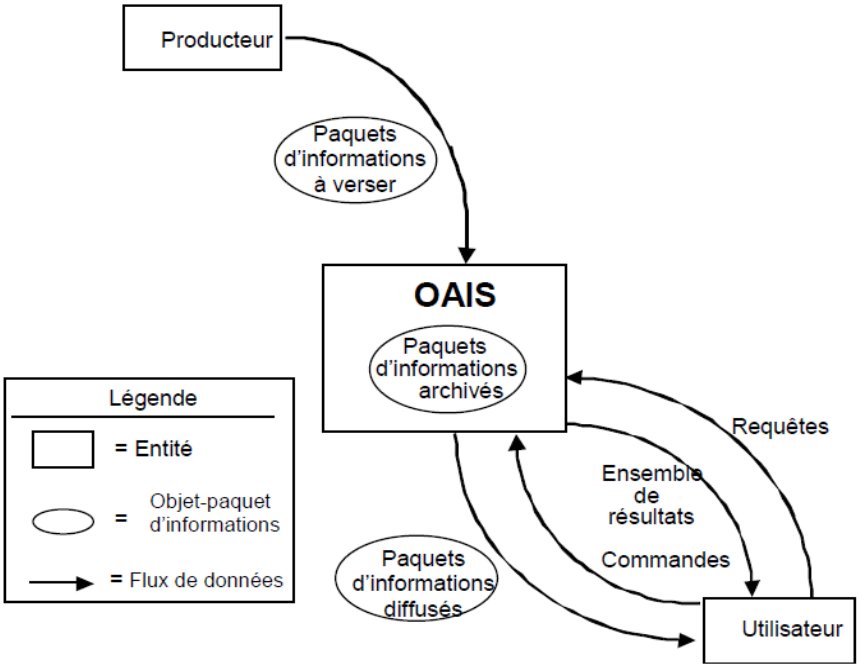
Afin de garantir ses différentes missions sur le long terme, le modèle de référence OAIS propose une architecture de référence qui consiste en une organisation de personnes et de systèmes ayant la responsabilité de préserver l'information pour une communauté déterminée d'un domaine spécifique (Brunsmann et al. 2012, p. 27). OAIS repose sur deux modèles : le modèle fonctionnel et le modèle d'information.

Le modèle fonctionnel définit un certain nombre de rôles, d'entités fonctionnelles et de responsabilités relatives, y compris les différentes relations qu'ils peuvent avoir entre eux (Banat-Berger, Duplouy, Huc 2009, p. 41, 52-55). Il propose en outre une terminologie permettant de garantir un niveau d'abstraction sur l'ensemble du modèle, et par là même de s'affranchir de toute application ou contexte professionnel, ce qui le rend adaptable à toutes les formes d'informations numériques (Ibid., p. 41). Ce niveau d'abstraction et cette émancipation de toute application ou contexte professionnel, garantit en outre la validité du modèle de référence en regard des évolutions technologiques (Ibid.).

Le modèle d'information, quant à lui, définit la composition des objets destinés à être conservés. Ce modèle s'appuie sur l'idée que l'information est organisée sous forme de

paquets d'information regroupant "plusieurs types d'informations à pérenniser afin de garantir une conservation sur le long terme" (Centre Informatique National de l'Enseignement Supérieur 2014). Ainsi, il y aura lieu de regrouper des informations sur les spécifications de formats, des informations de provenance et de contexte, des informations d'identités, voir encore des informations d'intégrité (Ibid.). Ces paquets d'information sont de nature différente selon leur position face à l'*Archive* – soit l'organisation chargée de conserver l'information (Consultative Committee for Space Data Systems 2012, p. 1-9 ; Roussel 2012-2014, p. 64). Trois types de paquets peuvent ainsi être définis : les *Paquets d'information à verser* (SIP) qui sont reçus par l'*Archive*, les *Paquets d'information à archiver* (AIP) qui sont conservés par l'*Archive*, et les *Paquets d'information à diffuser* (DIP) qui sont diffusés par l'*Archive*.

Flux de données et paquets d'information



(Comité consultatif pour les systèmes de données spatiales 2005)

Le modèle de référence OAIS a servi de base pour le développement de nombreuses normes, et est aujourd'hui adopté par un grand nombre de communautés de

professionnels (Roussel, 2012-2014, p. 60, 67). De ce fait, il est l'élément incontournable pour la pérennisation de l'information électronique.

3.3.2.2 Principes et stratégies de pérennisation

Pour garantir la pérennité de l'information, différents principes et stratégies peuvent être mis en place. Comme vue précédemment, l'obsolescence est la cause principale de la perte d'information. Que ce soit au niveau du matériel, du logiciel, ou encore du format de représentation, les risques sont d'autant plus grands que les dispositifs de lecture et l'organisation de l'information, sont complexes. L'implémentation d'un système d'archivage électronique à long terme sur le modèle de référence OAIS doit notamment intégrer et détailler les stratégies en matière de planification de la pérennisation comme la migration, l'émulation, la virtualisation, ou encore la transformation (Brunsmann et al. 2012, p. 28, 29). Bien que des différences subsistent, un regroupement peut être effectué entre certaines de ces stratégies. Ainsi, émulation et virtualisation peuvent être considérées comme une stratégie en soit, alors que migration et transformation, comme une autre.

"Emulation – combines software and hardware to reproduce running of another computer with all its main characteristics, allowing programs or media designed for certain environment to be executed in different, most often newer environments" (Deljanin 2012, p. 51). Cette stratégie de pérennisation, qui consiste à simuler le fonctionnement d'un logiciel, prendrait donc tout son sens dans l'optique de la disparition d'un élément logiciel du dispositif de lecture. Cependant, l'absence d'acteurs sur le marché en faveur de cette technique, combinée à différentes contraintes techniques[13], représente un certain nombre de risques tant techniques qu'économiques (Banat-Berger, Duplouy, Huc 2009, p. 124).

Second regroupement de stratégies, la migration, l'une des stratégies les plus couramment appliquées pour répondre à l'obsolescence matériel et logiciel (Harvey 2010, p. 165), peut être caractérisée comme suit : "Migration refers to transferring digital materials from one technology (one generation of hardware or software) to another, or from one format to another, preserving their intellectual content and retaining the ability for users to retrieve, display, and otherwise use them" (Ibid., p. 168). Différents types de migrations peuvent être mis en exergue :

- **le rafraichissement** : le *Contenu d'information* est transféré sur un support du même type. Principalement utilisé pour des raisons de dégradation du support,

[13] La principale contrainte est le besoin de disposer d'un nombre important d'émulateurs spécifiques. La conception d'émulateurs fonctionnant avec un grand nombre de type de matériel étant difficile à réaliser (Banat-Berger, Duplouy, Huc 2009, p. 124).

constatée ou planifiée (Banat-Berger, Duplouy, Huc 2009, p. 120-121 ; Consultative Committee for Space Data Systems 2012, p. 5-4, 5-5) ;

- **la duplication** : le *Contenu d'information* est transféré sur un support d'un nouveau type. Notamment utilisée pour des raisons de besoins en terme de volumétrie, d'attentes des utilisateurs, ou encore lorsque la technologie du support se trouve en fin de vie (Banat-Berger, Duplouy, Huc 2009, p. 121 ; Consultative Committee for Space Data Systems 2012, p. 5-4 - 5-6) ;
- **le ré-empaquetage** : le *Contenu d'information* est transféré sur un support d'un nouveau type, avec une modification de l'organisation des fichiers et des répertoires. Ce ré-empaquetage peut être la conséquence de la duplication (Banat-Berger, Duplouy, Huc 2009, p. 121 ; Consultative Committee for Space Data Systems 2012, p. 5-5, 5-6) ;
- **la transformation** ou **migration de formats** : le *Contenu d'information*, au sens du modèle OAIS, est modifié. Il y a modification de l'organisation de la séquence de bits. Cette transformation est utilisée pour répondre à l'indisponibilité de matériels et/ou logiciels du fait d'une obsolescence technologique, à l'évolution des besoins des utilisateurs, ou encore à des impératifs en termes de pérennisation des *Contenus d'informations* (Banat-Berger, Duplouy, Huc 2009, p. 122 ; Consultative Committee for Space Data Systems 2012, p. 5-5 - 5-7).

Les impératifs en termes de pérennisation, spécifiés dans le cas de la migration de formats, peuvent reposer entre autre sur un certain nombre de formats, considérés comme "acceptables pour la pérennisation" (Banat-Berger, Duplouy, Huc 2009, p. 122). Quelques principes peuvent ainsi être mis en évidence concernant les formats de représentation de l'information aptes à la conservation à long terme. Trois principes semblent suffisamment importants pour être rappelés ici : ouverture, indépendance et diffusion.

Un format est dit ouvert si sa documentation est disponible et accessible à tous. L'ouverture d'un format facilitera la création d'outils de restitution ou de conversion vers un nouveau format (Banat-Berger, Duplouy, Huc 2009, p. 105). Cette mise à disposition de la documentation peut se faire selon différentes modalités. Ainsi, une ouverture du format n'implique pas forcément la gratuité de la mise à disposition de la documentation (Centre de coordination pour l'archivage à long terme des documents électroniques 2014).

L'indépendance d'un format vis-à-vis de son fournisseur est la garantie d'une certaine stabilité en regard d'impératifs économiques ou stratégiques (Centre de coordination pour l'archivage à long terme des documents électroniques). Cette indépendance ne doit pas se faire uniquement en regard de son fournisseur, mais devrait aussi être effective envers d'autres éléments : formats, systèmes d'exploitation, matériels (Banat-Berger, Duplouy, Huc 2009, p. 106). Par exemple, certains logiciels peuvent être dépendants de systèmes de lecture déterminés.

Le critère de la diffusion est un critère non négligeable. En effet, "une large diffusion d'un format, c'est-à-dire une forte quantité de données existantes dans un tel format, contribue à sa grande longévité, étant donné que le nombre considérable de fichiers à convertir s'oppose à un changement rapide de format" (Centre de coordination pour l'archivage à long terme des documents électroniques 2014). Une large diffusion d'un format présuppose aussi l'existence d'un nombre suffisamment important d'applications et d'outils compatibles avec ce dernier.

Différentes organisations proposent des registres de formats de représentation. Ces derniers sont une aide en termes de pérennisation. Ils décrivent les formats numériques et leurs spécifications et permettent une normalisation des métadonnées administratives renseignant le format des objets numériques. Parmi ces registres citons PRONOM[14] des Archives nationales anglaises (Banat-Berger, Duplouy, Huc 2009, p. 108, 109).

3.3.3 Records Continuum et archives électroniques

Le modèle de cycle de vie des documents communément admis dans le monde de l'archivistique est celui de la théorie dite des trois âges des archives (Caya 2004). Ce dernier repose sur une succession de périodes, de la création au sort final, où l'activité du document décroît (actif, semi-actif, inactif) et où sa valeur se modifie (valeur primaire et secondaire) (Kent 2002, p. 68). Si certaines critiques sont venues par le passé remettre en question ce modèle, l'électronique a permis l'émergence d'une nouvelle représentation du cycle de vie des documents : le Records Continuum (Ibid., p. 69).

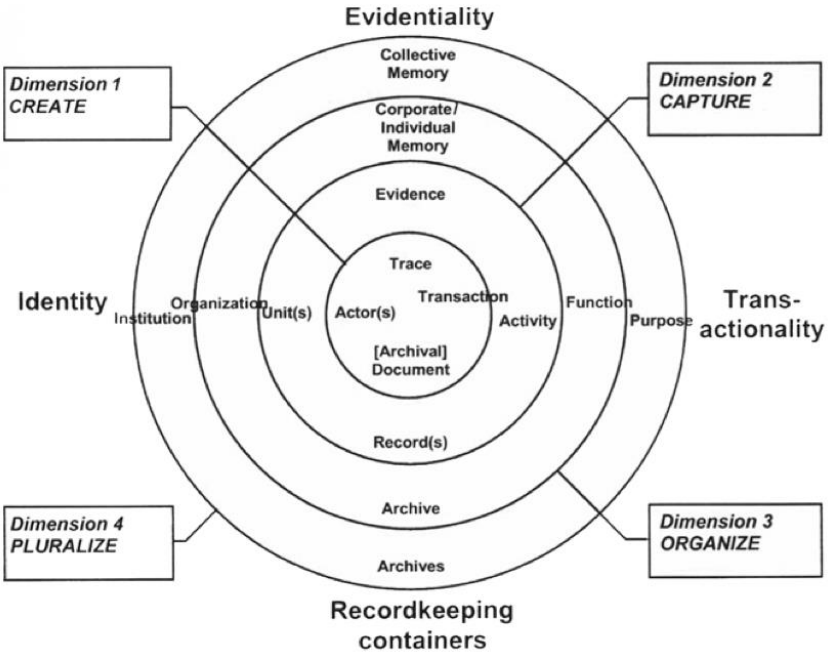
Contrairement à la vision "classique" du cycle de vie des documents basé sur une successivité d'étape, le modèle du Records Continuum s'en affranchit et propose une vision où le document d'archives est pris dans sa continuité. "The Records Continuum model provides a concept of service to users troughout the lifetime of records. Its standards-compliant metadata model, applied once in a specific environment, can then be used many times to meet a range of business purpose" (An 2005, p. 64).

Le succès de ce modèle dans l'environnement électronique tient au fait qu'il est difficile de séparer les documents d'archives en plusieurs âges ou périodes, parce que les étapes traversées par les documents d'archives peuvent être récurrentes et concerner toutes les périodes (Svärd 2013, p. 165). Ce modèle insiste sur la nature probante, transactionnelle et contextuelle du document d'archives et apporte une vue multidimensionnelle de sa création dans son aspect social et organisationnel (Ibid.).

[14] www.nationalarchives.gov.uk/PRONOM/ [Consulté le 11.02.2015].

Le schéma plus loin présente les quatre dimensions du modèle du Records Continuum, soit *Create*, *Capture*, *Organize*, et *Pluralize* (Flynn 2001, p. 82-83) :

Le modèle du Records Continuum



(http://exegesis.blackaeonium.net/chapter1.html)

La première dimension, *Create*, implique la présence d'un ou plusieurs créateurs (*Actor(s)*) dont la transaction (*Transaction*) génère un document (*Document*) qui en est la trace (*Trace*).

Deuxième dimension, *Capture*, implique l'unité organisationnelle (*Unit(s)*) à laquelle le créateur est associé, l'activité (*Activity*) dans le contexte de laquelle a lieu la transaction, le document créé avec ses informations contextuelles (*Records*), ainsi que la preuve de l'activité (*Evidence*).

Dans la troisième dimension, *Organize*, l'organisation (*Organisation*) est liée à ses fonctions (*Function*) et aux activités (*Activity*) qui les constituent, aux Archives en tant que somme des documents d'une organisation (*Archive*), ainsi qu'à sa propre mémoire (*Corporate / Individual Memory*).

Enfin, quatrième dimension du modèle, *Pluralize*, représente la place des Archives en tant que somme des documents de plusieurs organisations (*Archives*). Ces derniers

représentent la mémoire collective (*Collective Memory*). Le terme institution (*Institution*) est utilisé pour refléter la reconnaissance sociale des organisations, alors que les objectifs sont assimilés aux fonctions envisagées dans une perspective sociétale plus large (*Purpose*).

3.3.4 Conclusion sur l'archivage électronique à long terme

L'omniprésence des technologies numériques dans la vie de tous les jours de façon générale, et dans les organisations de façon plus particulière, participe à l'émergence de nouveaux défis. Gérer des documents électroniques, que ce soit pour des raisons patrimoniales ou pour des raisons légales ou d'information, posent des obstacles en termes de pérennisation. En effet, l'obsolescence des dispositifs de lecture capables de restituer l'information, peut amener le risque de rendre indisponible tout un pan de l'information produite.

Au-delà de garantir l'accessibilité à leurs informations, les organisations doivent aussi s'assurer de préserver leur valeur probante. Garantir l'authenticité, la fiabilité et l'intégrité des archives électroniques, s'assurer que ces dernières ne subissent pas de dégradations, et finalement planifier leur pérennisation sont les trois missions principales d'un système d'archivage électronique.

Afin de répondre aux différentes missions citées plus haut, la norme OAIS propose un modèle fonctionnel et un modèle d'information devant permettre la mise en place d'une architecture de référence à même de préserver l'information pour une communauté spécifique. A côté de ces modèles, diverses stratégies et principes de pérennisation peuvent être mis en place. Si certaines stratégies se relèvent impossible à mettre en œuvre, à l'instar de la préservation des technologies de restitution, d'autres, comme la migration, offrent de bonnes perspectives en termes d'implémentation.

C'est au travers de la documentation que ces stratégies et ces principes peuvent répondre à la problématique de la pérennisation. En effet, disposer d'information sur les spécifications techniques des objets numériques à conserver, sur leur provenance et leur contexte de création, voire sur leur intégrité contribue grandement à garantir leur conservation et l'exploitation de leurs caractères probatoires et informationnels.

A cette fin, les normes et standards sont des aides non négligeables en termes de documentation. Leur utilisation est donc une plus-value pour l'archivage électronique à long terme.

4. Analyse des normes et standards

Ce chapitre a pour objet l'analyse proprement dite des normes et standards en matière d'archivage électronique à long terme, et l'identification des éléments participant à la mesure des dimensions de la qualité des archives électroniques. En préambule, un bref aperçu choisi sera donné des normes archivistiques de façon générale, et plus particulièrement de celles pouvant amener des éléments de mesure de la qualité des archives. Parmi ces dernières, celles se traduisant par l'utilisation de métadonnées sont particulièrement enclin à fournir des éléments de mesure de qualité. Cet aperçu a notamment pour objectif de visualiser le périmètre normatif utilisé pour ce travail. Soit de comprendre les raisons qui ont amené la sélection de certains textes au détriment d'autres. Il doit aussi ouvrir des perspectives futures en matière de mesure de la qualité des archives. En effet, si le choix a été de s'orienter sur les normes en matière d'archivage électronique à long terme pour des raisons de volume de travail, certaines normes semblent indispensables pour une mesure globale et complète de l'ensemble des dimensions de la qualité des archives. Pour finir, et avant l'analyse proprement dite, les textes normatifs retenus seront présentés et leur choix justifiés.

4.1 Aperçu sélectif des normes archivistiques

L'univers normatif dans le domaine des archives est relativement vaste et complexe. En dehors des normes générales dites de système de management qui intéressent plusieurs aspects d'une organisation (Arnaud 2012-2014, p. 112)[15], de nombreux apports normatifs concernent tout particulièrement voire exclusivement les archives. L'organisation de ces normes et standards peut se faire selon différentes approches : organisme d'élaboration, fonction archivistique associée, ou encore selon leur place dans le cycle de vie des documents comme le montre la figure située plus loin.

Trois étapes de ce cycle de vie des documents semblent tout particulièrement importantes afin de disposer d'un panorama normatif suffisant pour le présent travail : *CAPTURE*, *PRESERVE*, *MANAGE*.

[15] Management de la qualité (ISO 9000 et 9001), Management environnemental (ISO 14001), Management du risque (ISO 31000) et Management de la sécurité de l'information (ISO 27000 et 27001).

Étape du cycle de vie : *CREATE – PRESERVE - DESTROY*

(www.patrimoine.vd.ch/fileadmin/groups/19/PDF/AFNOR-CN11_bilingue.pdf)

4.1.1 Les normes de description archivistique

La première étape de ce cycle de vie, la capture (*CAPTURE*) ou saisie – parfois intégrée à l'étape de création – comprend les normes dites de description. Dans cette famille se trouvent notamment les quatre normes publiées par le Conseil International des Archives (CIA)[16] que sont ISAD(G), ISAAR, ISDF et ISDIAH. La *Norme générale et internationale de description archivistique* ISAD(G) est la norme pour la description de fonds d'archives. La description archivistique a pour but "d'identifier et d'expliquer le contexte et le contenu des documents d'archives, en vue de faciliter leur accès" (Conseil international des archives 2000, p. 7). A cette fin, 26 éléments d'information spécifiques pouvant être combinés entre eux sont renseignés, dans l'optique de créer une description d'une entité archivistique. Cette description peut s'appliquer à

[16] http://www.ica.org/ [Consulté le 02.01.2015].

l'ensemble des unités de description, quelle que soit leur forme ou leur support. (Ibid.). ISAD(G) énonce quatre règles fondamentales pour la description à plusieurs niveaux. Tout d'abord, la description doit se faire du général au particulier, c'est-à-dire que les informations doivent coïncider au niveau de description correspondant (fonds, sous-fonds, série, dossier…). Deuxièmement, elles doivent être adaptées au dit niveau de la description. Dans ce sens, seul les informations adéquates doivent être données au niveau décrit. Troisièmement, un lien doit permettre de situer l'unité de description dans la hiérarchie. Et enfin, il y a lieu de ne pas répéter les informations déjà contenues dans les niveaux de description hiérarchiquement supérieurs. (Ibid., p. 13). Des points d'accès peuvent être établis entre l'élément d'information *producteur* de la notice descriptive ISAD(G) et la norme ISAAR pour faciliter la recherche.

La *Norme Internationale sur les notices d'autorité utilisées pour les Archives relatives aux collectivités, aux personnes ou aux familles* ISAAR (CPF) "fournit les lignes directrices pour la préparation de notices d'autorité qui offrent la description d'entités (collectivités, personnes ou familles) associées à la production et à la gestion des archives" (Conseil international des archives 2004, p. 7). Son objectif est donc de décrire le contexte de création à travers la description des producteurs d'archives, qu'ils soient des collectivités, des personnes ou des familles. A cette fin, 31 éléments répartis dans cinq zones peuvent être renseignés. La raison d'utiliser une norme spécifique pour la description des notices d'autorité tient du fait de pouvoir disposer de notices indépendantes afin de les associer à des descriptions issues de plusieurs institutions. Ces notices permettent en outre d'avoir une information contextuelle normalisée et donc échangeable (Ibid.). Comme indiqué plus haut, la notice d'autorité peut être reliée à la zone du contexte de la norme ISAD(G).

La *Norme internationale pour la description des fonctions* ISDF "fournit les lignes directrices pour la description des fonctions des collectivités associées à la production et à la gestion des archives" (Conseil international des archives 2007, p. 7). Alors qu'ISAAR s'occupe du contexte lié au producteur, ISDF a pour objectif de décrire le contexte de création à travers la description des fonctions. A l'aide de 23 éléments répartis en quatre zones, ISDF tente d'expliquer le pourquoi et le comment de la création des documents d'archives (Ibid.). Cette norme a un intérêt tout particulier pour la description du contexte, car, "en général, les fonctions sont reconnues comme plus stables que les structures administratives qui sont souvent fusionnées ou transférées lors de réorganisations" (Ibid.). Cette norme est destinée à compléter les descriptions archivistiques et les notices d'autorité.

La *Norme internationale pour la description des institutions de conservation des archives* ISDIAH "donne des règles générales pour la normalisation de descriptions d'institutions conservant des fonds d'archives" (Conseil international des archives 2008, p. 9). Il s'agit encore ici d'une norme permettant de comprendre le contexte, mais cette fois ci du point de vue de la conservation des documents. "Les détenteurs de documents d'archives peuvent être décrits dans des notices d'autorité conformes à ISAAR(CPF) comprenant les éléments de description appropriés indiqués dans ISDIAH" (Ibid.) Autrement dit, ISDIAH peut être interprétée comme une sous-catégorie d'ISAAR(CPF).

Parmi les quatre normes de description citées plus haut, deux possèdent actuellement un équivalent en format d'encodage XML. EAD, pour *Encoded Archival Description*, est un standard d'encodage pour instruments de recherche archivistique. L'un de ses principes directeurs est de maintenir sa compatibilité avec ISAD(G) (Society of American Archivists and the network development and Marc standard office of the Library of Congress 2002, p. V). Un autre standard d'encodage, EAC, pour *Encoded Archival Context*, va de pair avec la norme ISAAR(CPF) : "EAC-CPF is based on the International Standard Archival Authority Record for Corporate Bodies, Persons and Families […]. It is used closely in association with Encoded Archival Description (EAD), an XML-Schema for encoding archival finding aids, but not limited to it" (Society of American Archivists, Staatsbibliothek zu Berlin). En ce qui concerne les autres normes, aucun format d'encodage n'existe officiellement, bien qu'un format dédié à ISDF ait été proposé dernièrement[17] (Dubois, Wildi 2013). Il est particulièrement intéressant, dans le cadre de ce travail, de pouvoir utiliser des formats d'encodage, car cela ouvre des perspectives futures pour la mesure de la qualité des archives et son automatisation.

4.1.2 Les normes de préservation

Seconde étape de notre cycle de vie des documents, *PRESERVE*, a pour objet les normes dites de préservation. Si ces dernières doivent couvrir l'ensemble du spectre des archives et leurs supports – analogiques et électroniques – l'accent sera mis ici uniquement sur le volet électronique, et plus particulièrement sur les profils de métadonnées. Ces derniers sont en effet capitaux pour mesurer les dimensions de la qualité des archives.

Les normes et standards ont un rôle central à jouer dans le domaine de la pérennisation des contenus électroniques. Ils permettent tout d'abord de disposer d'un

[17] Ce format proto-EAF a été proposé par les Archives de l'État du Valais et la société Docuteam GmbH.

"fondement stable et relativement pérenne dans un environnement technologique qui ne cesse de se transformer" (Roussel 2012-2014, p. 59). Ils offrent ensuite des concepts et une terminologie permettant la compréhension et le dialogue entre différentes communautés, et par la même contribuent à l'élaboration de méthodologies et d'outils logiciels (Banat-Berger, Duplouy, Huc 2009, p. 36). Enfin, leur utilisation et leur adaptation sont un facteur de réduction des coûts de mise en œuvre (Wildi 2012, p. 14).

Si la norme centrale dans le domaine est sans conteste la norme ISO 14721, plus connue sous le nom OAIS, pour *Open Archival Information System*, différentes initiatives ont vu le jour dans son sillage (Hofman 2005, p. 27). C'est le cas du standard d'encodage et de transfert de métadonnées METS, ou encore du dictionnaire de métadonnées de préservation d'objets numériques PREMIS (Library of Congress 2011 ; PREMIS Editorial Committee 2012, p. 3).

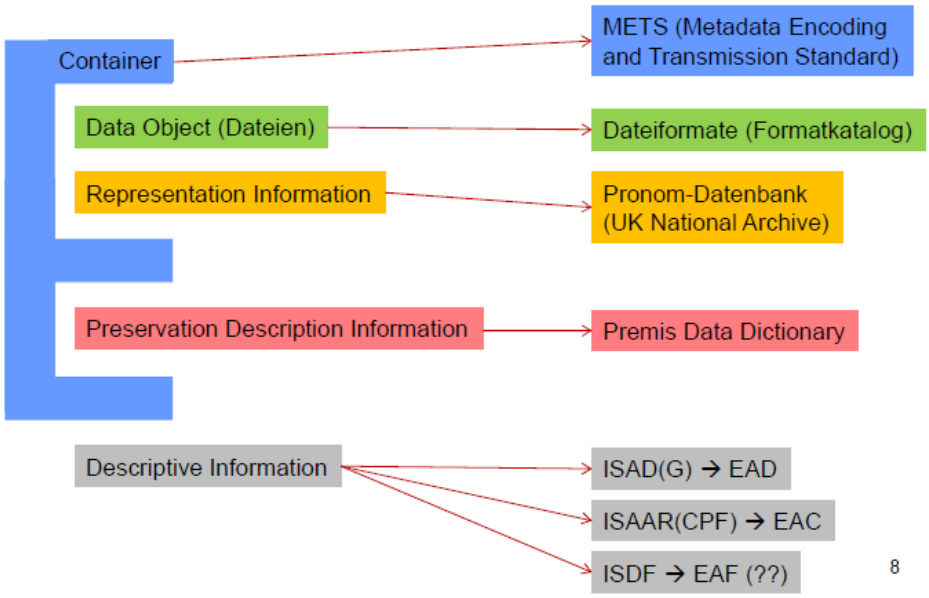
D'autres contributions se fondent également sur le modèle de référence OAIS. C'est le cas notamment des textes liés aux processus de versement PAIMAS et PAIS. La norme PAIMAS pour *Producer Archive Interface Methodology Abstract Standard* "vise à définir, pour l'archive et le producteur, une méthodologie de travail commune et structurée" dans le cadre d'un processus de versement (Banat-Berger, Duplouy, Huc 2009, p. 57). PAIMAS s'intéresse de ce fait à l'entité *Entrées* du modèle de référence OAIS (Consultative Committee for Space Data Systems, 2004, p. 1-1). De son côté, le projet de norme *Producer-Archive Interface Specificatio*n (PAIS), propose quant à lui une représentation formelle des objets versés à l'*Archive* au travers des SIP (Banat-Berger, Duplouy, Huc 2009, p. 60). L'univers normatif dans le domaine de la préservation numérique est donc relativement riche.

En s'intéressant aux profils de métadonnées et aux différentes combinaisons que ces dernières peuvent engendrer pour décrire les objets numériques à pérenniser, la richesse de cet univers normatif prend une dimension supplémentaire. Un profil de métadonnées peut être défini comme "les différentes métadonnées employées et regroupées en catégories ainsi que leurs relations" (Kalbermatten 2012, p. 34). Les profils de métadonnées sont importants, car ils permettent de sélectionner les métadonnées utiles à une organisation pour pérenniser et comprendre ses données électroniques. Actuellement, deux profils sont utilisés en Suisse : le *Matterhorn METS Profile* et le standard *eCH-0160*. Tous les deux ont fait l'objet d'une standardisation (Kalbermatten 2012, p. 34, 35, 40).

Le standard *eCH-0160*, issu d'un profil utilisé aux Archives fédérales suisses, est admis comme standard national par l'association eCH[18] depuis fin novembre 2012 (Kalbermatten 2012, p. 35, 36). Il repose sur un dictionnaire de données et un schéma XML et modélise un *Paquet d'information à verser*. Il ne s'agit donc pas d'un modèle utilisable pour une conservation à long terme. Le standard *eCH-0160* nécessite donc d'être transféré en un format d'archives, soit être transformé en *Paquet d'information à archiver* (Kalbermatten 2012, p. 36 ; Nef, Wildi 2013).

Contrairement au profil *eCH-0160*, le *Matterhorn METS Profile* repose sur des standards internationaux et couvre l'ensemble du modèle de référence OAIS (Nef, Wildi 2013). Elaboré conjointement par la société Docuteam et les Archives d'État du Valais, ce profil "s'appuie sur les deux standards internationaux et principaux de l'archivage numérique : METS et PREMIS" (Kalbermatten 2012, p. 40). En outre, il intègre les trois normes suivantes du Conseil international des archives : ISAD(G) pour la description du contenu, ISAAR(CPF) pour les producteurs, et ISDF pour les fonctions (Ibid.). Le *Matterhorn METS Profile* a été enregistré par la Library of Congress en novembre 2012 (Ibid.).

Le Matterhorn METS Profile et les standards de métadonnées archivistiques



(www.cmistar.ch/.../Tobias_Wildi_-_Live_OAIS.pdf)

[18] http://www.ech.ch/vechweb/page [Consulté le 09.02.2015].

Les normes et standards en général, et les profils de métadonnées en particulier sont centraux dans le monde des archives électroniques. Ce sont en effet des facteurs de stabilité dans un environnement en constante évolution, qui permettent le dialogue et la compréhension entre communautés diverses, et sont donc de formidables vecteurs de développement d'outils et de réduction des coûts. Par ailleurs, combiner des sets de métadonnées différents peut permettre de mesurer le spectre entier des dimensions de la qualité des archives.

4.1.3 Les normes sur les documents d'activité

MANAGE, dernière étape du schéma normatif, concentre notamment les normes ISO[19] sur les documents d'activité. Ces dernières font partie d'un ensemble plus vaste connu sous le nom de normes dites de systèmes de management. Parmi celles-ci, citons la norme ISO 15489 sur le Records Management, la norme ISO 30300 sur les Systèmes de gestion des documents d'activité, ou encore ISO 23081 sur les métadonnées pour la gestion des documents d'activité.

Ces normes constituent des soutiens "pour conceptualiser, implémenter, surveiller et améliorer les processus et contrôles sur les documents d'activité" (Cottin, Dessolin-Baumann 2012-2014, p. 123). Les documents d'activité se doivent d'intégrer des métadonnées. Ces dernières permettent de décrire "le contexte, le contenu et la structure des documents d'activité ainsi que leur gestion dans le temps" (Organisation internationale de normalisation 2011a, p. 8).

L'antériorité de la position de ces normes par rapport à d'autres fonctions archivistiques est importante pour ce travail, du fait que les exigences formulées pour les documents et les systèmes d'activité se retrouveront en partie sous formes de qualités des archives. Les exigences que ces normes sollicitent, notamment sous forme de métadonnées, sont donc à prendre en compte dans cette analyse. Dans un environnement électronique, du fait de la nature des documents, il y a lieu de penser à l'archivage dès la création de systèmes d'information (Dunant Gonzenbach, Ducry 2013 p. 1). La prise en compte des normes sur les documents d'activité est donc un impératif dans notre propos.

4.2 Choix des normes et standards analysés

La sélection des normes et standards à analyser dans le cadre de ce travail s'est portée sur trois textes normatifs : une norme générale et conceptuelle, le modèle de

[19] http://www.iso.org/iso/fr/ [Consulté le 09.02.2015].

référence OAIS, un standard d'encodage et de transfert de métadonnées, le schéma METS, et enfin, un dictionnaire de données pour les métadonnées de préservation d'objets numériques, PREMIS.

Le choix du modèle de référence OAIS s'est imposé facilement du fait de sa place centrale dans le domaine de la pérennisation de l'information et aussi du fait que la majeure partie des systèmes d'archivage électronique mis en œuvre actuellement s'appuie sur son cadre conceptuel (Roussel 2012-2014, p. 60). Pour ce qui est de PREMIS, il s'agit "d'un des dictionnaires de métadonnées les plus complets et les plus détaillés aujourd'hui" (Banat-Berger, Duplouy, Huc 2009, p. 56) "définissant des éléments de métadonnées jugés indispensables à une fonction d'archivage pérenne" (Ibid., p. 144). Qui plus est, l'activité de maintenance, soutenue par la Library of Congress[20], est très dynamique (Library of Congress Network Development and MARC Standards Office 2009, p. 4). METS fournit, quant à lui, "un mécanisme souple pour l'encodage des métadonnées descriptives, administratives et de structure d'un objet numérique, ainsi que pour exprimer les liens complexes entre ces différentes formes de métadonnées" (Library of Congress 2011). Il s'agit d'un format de type "capsule" ou "conteneur", basé sur le modèle de référence OAIS, qui peut intégrer d'autres formats de métadonnées descriptives ou techniques (Banat-Berger, Duplouy, Huc 2009, p. 137). Le couple METS – PREMIS[21] est couramment utilisé, que ce soit dans des institutions à vocation patrimoniale, comme dans le projet SPAR de la Bibliothèque Nationale de France (Bibliothèque nationale de France 2013) et le projet Planets supporté par la British Library (Preservation and Long-term Access through Networked Services 2007), ou encore pour des organisations de diverses natures, telles que les entreprises ABB et SRG-SSR, la ville de Baden, ou encore les Archives d'État de Bâle Campagne ou du Valais (Wildi 2012, p. 14). A l'instar de la dernière institution citée, l'association des standards METS et PREMIS est aussi l'occasion de créer des profils de métadonnées spécifiques. Le modèle de données *Matterhorn METS Profile*[22] en est un bon exemple. Développé par les Archives d'État du Valais et la société Docuteam dans le cadre d'un projet d'archivage du processus décisionnel du Conseil d'État valaisan, *Matterhorn METS Profile* est un profil conçu "pour verser, conserver et partager des paquets d'information provenant d'archives, de bibliothèques ou encore de musées" (Dubois, Wildi, 2012).

[20] www.loc.gov/standards/premis/ [Consulté le 03.11.2014].
[21] METS pour le format d'empaquetage et PREMIS pour les métadonnées de pérennisation.
[22] Le *Matterhorn METS Profile* est basé sur l'EAD et PREMIS pour encoder respectivement les métadonnées descriptives et les métadonnées techniques et administratives, alors que METS sert de format d'empaquetage.

4.3 OAIS

Le premier texte normatif choisi pour cette analyse est la norme ISO 14721 *Systèmes de transfert des informations et données spatiales – système d'archivage d'information (SOAI)*[23], plus connu sous le nom de modèle de référence OAIS pour *Open Archival Information System*. Les termes spécifiques à la norme sont indiqués en italique et commencent par une majuscule. Lors de leur première apparition, ils sont suivis par leur équivalent anglais entre parenthèses.

4.3.1 Origine et développement

Les premières réflexions relatives à l'archivage électronique à long terme sont intimement liées au domaine aérospatial. Les deux principales agences de ce secteur, la NASA (National Aeronautics and Space Administration) et le CNES (Centre National d'Étude Spatiales) sont des producteurs de données électroniques depuis près d'un demi-siècle. Ces données sont conservées au sein même de ces institutions, ou alors dans un service dédié (Banat-Berger, Duplouy, Huc 2009, p. 39). Dans ce contexte, ces organisations sont à même d'entamer des réflexions dans le domaine de l'archivage électronique à long terme. C'est ainsi que le CCSDS (Consultative Committee for Space Data Systems), groupe de standardisation dont sont membres les différentes agences spatiales, a élaboré le modèle de référence OAIS (Banat-Berger, Duplouy, Huc, 2009, p. 35). Ce standard, devenu la norme ISO 14721:2003, a reçu l'apport de nombreux organismes externes au domaine aérospatial, du fait de sa portée qui dépasse largement cet univers. Ainsi, la Bibliothèque nationale de France (BnF) et la National Archives and Records Administration (NARA) ont contribué à son élaboration (Banat-Berger, Duplouy, Huc 2009, p. 40 ; Hofman 2005, p. 27).

4.3.2 Définition et principes

Le modèle de référence OAIS "propose une vue globale cohérente de la question de l'archivage numérique" (Banat-Berger, Duplouy, Huc 2009, p. 41). A cette fin, il définit un certain nombre de concepts et de fonctions et propose en outre un vocabulaire qui lui permet de s'affranchir de tout contexte, application ou domaine spécifique (Ibid.). Cette indépendance lui permet de conserver une vue abstraite de la question de l'archivage électronique.

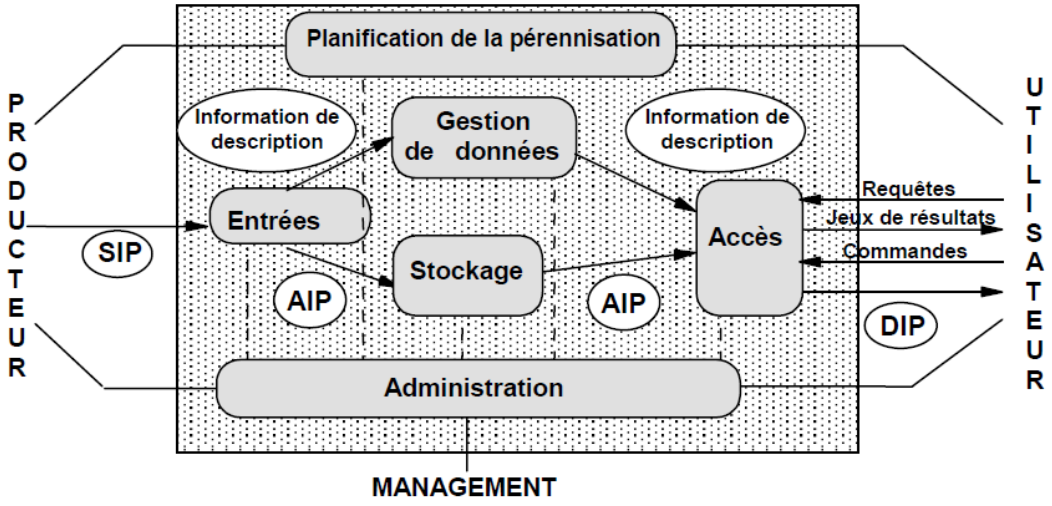
L'utilisation d'un vocabulaire spécifique est centrale dans l'environnement normatif en général et dans la norme OAIS en particulier. L'*Archive OAIS* (*Archive*) y est définie

[23] Pour des raisons de lisibilité, certaines citations sont issues de la version française de la norme OAIS datant de 2005. La conformité de ces dernières a été vérifiée en regard de la version anglaise datant de 2012.

comme "[l']organisation chargée de conserver l'information pour permettre à une Communauté d'utilisateurs cibles d'y accéder et de l'utiliser" (Comité consultatif pour les systèmes de données spatiales 2005, p. 1-8). Six responsabilités incombent à l'*Archive* afin de mener à bien sa mission (Roussel 2012-2014, p. 62, 63) :

- s'assurer que l'information qu'elle va recevoir correspond à sa mission et au besoin des utilisateurs ;
- garantir la pérennisation de l'information conservée grâce à une connaissance suffisante de cette dernière ;
- disposer d'une information de description suffisante pour permettre aux utilisateurs sa recherche ;
- déterminer quels sont les utilisateurs et s'assurer que l'information est compréhensible par ces derniers ;
- garantir la conservation de l'information, et permettre sa diffusion authentique à l'aide de stratégies et de politiques ;
- garantir la disponibilité de l'information conservée pour les utilisateurs cible.

Les entités fonctionnelles du modèle OAIS



(Comité consultatif pour les systèmes de données spatiales 2005)

OAIS repose sur deux modèles : un modèle fonctionnel et un modèle d'information. Le modèle fonctionnel, présenté dans la figure ci-dessus, recouvre toutes les activités indispensables d'une *Archive*. Quatre entités fonctionnelles de base y sont représentées : *Entrées* (*Ingest*), *Stockage* (*Archival Storage*), *Gestion de données* (*Data Management*), et enfin *Accès* (*Access*). A côté de ces entités de base, deux autres entités, *Planification de la pérennisation* (*Preservation Planning*) et

Administration (*Administration*), ont un rôle plus général sur l'*Archive* (Banat-Berger, Duplouy, Huc 2009, p. 52, 54).

L'entité *Administration* vise à assurer "une supervision continue du fonctionnement des autres entités de l'OAIS" (Comité consultatif pour les systèmes de données spatiales 2005, p. 1-9). Elle est en lien avec le *Management* (*Management*) qui définit la direction générale de l'*Archive* (Banat-Berger, Duplouy, Huc 2009, p. 44). L'entité *Planification de la pérennisation* a le rôle d'élaborer les stratégies en matière de conservation à long terme. A cette fin, elle effectue une triple surveillance, de la technologie, du *Producteur* (*Producer*), et de l'*Utilisateur* (*Consumer*) (Roussel 2012-2014, p. 67).

L'entité *Entrée* reçoit les objets numériques par le *Producteur* sous forme de *Paquets d'information à verser* (*Submission Information Package*) (SIP) et les transforme en *Paquets d'information archivés* (*Archival Information Package*) (AIP) une fois leur validation effectuée[24]. Ces paquets, transmis à l'entité *Stockage*, seront l'information que l'*Archive* devra conserver. Les *Informations de description*, éventuellement complétées, issues du *Producteur*, seront extraites et envoyées à l'entité *Gestion de données* (Banat-Berger, Duplouy, Huc 2009, p. 53)

Gestion de données, la dernière entité du modèle fonctionnel, maintient à jour les *Informations de description* (*Descriptive Information*) et les fournit à l'*Utilisateur* via l'entité *Accès*. *Accès* fait fonction d'interface pour fournir les objets numériques via des requêtes sur les *Informations de description*. Une fois la requête validée en fonction des droits des utilisateurs, l'entité *Accès* génère des *Paquets d'information diffusés* (*Dissemination Information Packages*) (DIP) sur la base des AIP conservés dans l'entité *Stockage* (Banat-Berger, Duplouy, Huc 2009, p. 53, 54).

Second modèle sur lequel repose OAIS, le modèle d'information part du constat qu'il n'est pas suffisant de conserver l'objet numérique, mais qu'un certain nombre d'informations doivent y être adjointes afin de le rendre compréhensible. Ainsi, les paquets d'information vus plus haut sont composés de trois types d'information : le *Contenu d'information* (*Content Information*), l'*Information d'empaquetage* (*Packaging Information*) qui lie les différents composants du paquet d'information, et les métadonnées requises pour lire et comprendre le *Contenu d'information* (Roussel 2012-2014, p. 64 ; Harvey 2010, p. 40 ; Kalbermatten 2012, p. 35).

[24] Cette validation se fait en fonction d'exigences fournies dans un *Protocole de versement*.

Parmi les métadonnées requises pour lire et comprendre le *Contenu d'information*, deux sont indispensables : l'*Information de représentation* (*Representation Information*) et l'*Information de pérennisation*[25] (*Preservation Description Information*) (PDI) (Banat-Berger, Duplouy, Huc 2009, p. 45, 49).

L'*Information de représentation* permet "de passer des bits constituant l'objet numérique au contenu informationnel de cet objet" (Banat-Berger, Duplouy, Huc 2009, p. 45). Pour exemple, une page web incluant des graphiques a besoin de la description de son environnement web (navigateur, etc.), du texte (codage de caractères ASCII) et du format de l'image pour être compréhensible (Harvey 2010, p. 41). Cette *Information de représentation* fait particulièrement appel aux normes décrivant les formats de représentation de l'information ou formats de données, mais inclus aussi des informations liés au codage des caractères (*Information de structure*) ou encore à la langue (*Information sémantique*) (Banat-Berger, Duplouy, Huc 2009, p. 47 ; Roussel 2012-2014, p. 64).

L'*Information de représentation* ne suffit pas pour rendre compréhensible le *Contenu d'information*. Une autre information est requise pour permettre à l'*Archive* d'assurer "l'ensemble de ses responsabilités" : l'*Information de pérennisation* (Banat-Berger, Duplouy, Huc 2009, p. 49). Cette dernière est constituée de quatre composants spécifiques :

- ***Information d'identification*** (*Reference Information*) : identifiant unique et permanent pour identifier et localiser le *Contenu d'information* ;
- ***Information de provenance*** (*Provenance Information*) : l'historique du *Contenu d'information* archivé ;
- ***Information de contexte*** (*Context Information*) : informations sur les relations du *Contenu d'information* avec d'autres objets ;
- ***Information d'intégrité*** (*Fixity Information*) : mécanismes permettant de démontrer l'authenticité du *Contenu d'information*, comme une empreinte numérique ou une somme de contrôle (Harvey 2010, p. 41).

Il faut ici rappeler que la norme OAIS est un modèle conceptuel qui identifie les catégories de métadonnées requises, "mais sans indiquer leur contenu exacte" (Kalbermatten 2012, p. 34). Il est donc indispensable de recourir à d'autres normes ou standards basés sur ce modèle pour implémenter une solution.

[25] Un troisième type de métadonnées" utile pour la compréhension du *Contenu d'information* se trouve sous la forme de l'information de description.

4.4 METS

Second texte normatif sélectionné, METS pour *Metadata Encoding and Transmission Standard* est un format XML d'encodage et de transfert de métadonnées utilisé notamment comme format conteneur. Les termes spécifiques au standard sont indiqués en italique. Lors de leur première apparition, ils sont suivis par leur équivalent anglais entre parenthèses.

4.4.1 Origine et développement

Le projet Making of America II (MOA2) est à l'origine de METS. Créé en 1997 par la Digital Library Federation, il avait pour objectif la création d'un format d'encodage de métadonnées descriptives, administratives et de structure, pour des objets numériques basés sur du texte et de l'image (Council on Library and Information Resources, 2014 ; Library of Congress 2011). METS, également créé par la Digital Library Federation, tente de développer ce projet initial et de produire un format d'encodage XML de métadonnées pour la gestion et l'échange d'objets numériques (Ibid.).

4.4.2 Définition et principes

Maintenu par la Library of Congress, "METS se présente sous la forme d'une structure XML permettant de transporter et de référencer des fichiers de données […] et des métadonnées associées" (Banat-Berger, Duplouy, Huc 2009, p. 64). Il détaille ainsi la structure de l'objet numérique, le nom et la localisation des fichiers le composant, ainsi que les métadonnées associées (Guenther, Xie 2007, p. 15). METS ne définit pas de métadonnées mais offre la possibilité de choisir ces dernières (Ibid.). Ces métadonnées peuvent être contenues dans le document METS, ou être référencées à l'extérieur du document (Library of Congress 2011). Dans ce sens, il peut servir d'unité de stockage (AIP), ou de transmission (SIP et DIP). METS est composé de sept sections principales (Dappert, Enders 2010, p. 10, 11 ; Library of Congress 2011, p. 1, 2) :

- ***en-tête METS*** (*METS Header*) : contient les métadonnées propres au document METS lui-même (création, modification…) ;
- ***métadonnées descriptives*** (*descriptive metadata*) : métadonnées de description (ex : EAD). Peuvent être encapsulées dans le fichier ou référencées à l'extérieur ;
- ***métadonnées administratives*** (*administrative metadata*) : informations sur la création des fichiers, leur provenance et les droits associés (ex : PREMIS). Peuvent être encapsulées dans le fichier ou référencées à l'extérieur ;
- ***section de fichiers*** (*file section*) : liste des fichiers. Peut-être regroupée par type de fichier (PDF, TIFF) ;

- ***carte de structure*** (*structural map*) : structure hiérarchique de l'objet numérique où tous les éléments sont reliés aux fichiers et aux métadonnées s'y rapportant. Seule section obligatoire dans METS ;
- ***liens structurels*** (*structural links*) : informations sur l'existence d'hyperliens entre éléments au sein de la carte de structure. Utilisés notamment pour archiver des sites web ;
- ***comportements*** (*behavior*) : associent des comportements exécutables au contenu de l'objet numérique.

METS fait fonction de conteneur de métadonnées pour la gestion et la transmission d'objets numériques. Dans ce sens, il peut servir de norme pour échanger des objets numériques entre différents entrepôts numériques (Library of Congress 2011). Comme indiqué auparavant, de nombreux projets et institutions l'associent à PREMIS pour les métadonnées administratives.

4.5 PREMIS

Dernier texte normatif sélectionné, PREMIS, qui signifie *PREservation Metadata Implementation Strategies*, désigne un dictionnaire dont l'objectif est de définir "des éléments de métadonnées jugés indispensables à une fonction d'archivage pérenne" (Banat-Berger, Duplouy, Huc 2009, p. 144), et qui s'intéresse ainsi à ce que l'on peut appeler les métadonnées de préservation. Les termes spécifiques au standard sont indiqués en italique. Lors de leur première apparition, ils sont suivis par leur équivalent anglais entre parenthèses.

4.5.1 Origine et développement

PREMIS doit son origine au groupe de travail éponyme sponsorisé conjointement par l'OCLC[26] (Online Computer Library Center) et le RLG[27] (Research Libraries Group). Ce groupe de travail composé de plus de trente membres issus de cinq pays et de plusieurs domaines d'activité comprenant des bibliothèques, musées, archives, agences gouvernementales, ou encore des acteurs du secteur privé, est en charge depuis plusieurs années d'un dictionnaire de données pour les métadonnées de préservation (PREMIS Editorial Committee 2012, p. 1, 2). La Library of Congress maintient le schéma XML pour PREMIS (Library of Congress Network Development and Marc Standards Office 2009, p. 4).

4.5.2 Définition et principes

PREMIS, tout comme METS, se base sur le modèle OAIS. Il peut être vu comme une traduction de ce modèle en *unités sémantique* (*semantic units*) implémentables

[26] https://oclc.org/home/ [Consulté le 06.11.2014].
[27] Le RLG et l'OCLC ont fusionnés en une entité commune le 1er juillet 2006. Source : www.oclc.org/research/partnership/history/ [Consulté le 06.11.2014].

(PREMIS Editorial Committee 2012, p. 3). A travers ses métadonnées, il cherche à s'assurer de l'intégrité, de la lisibilité et de la traçabilité des interventions des objets numériques (Library of Congress Network Development and Marc Standards Office 2009, p. 3, 4).

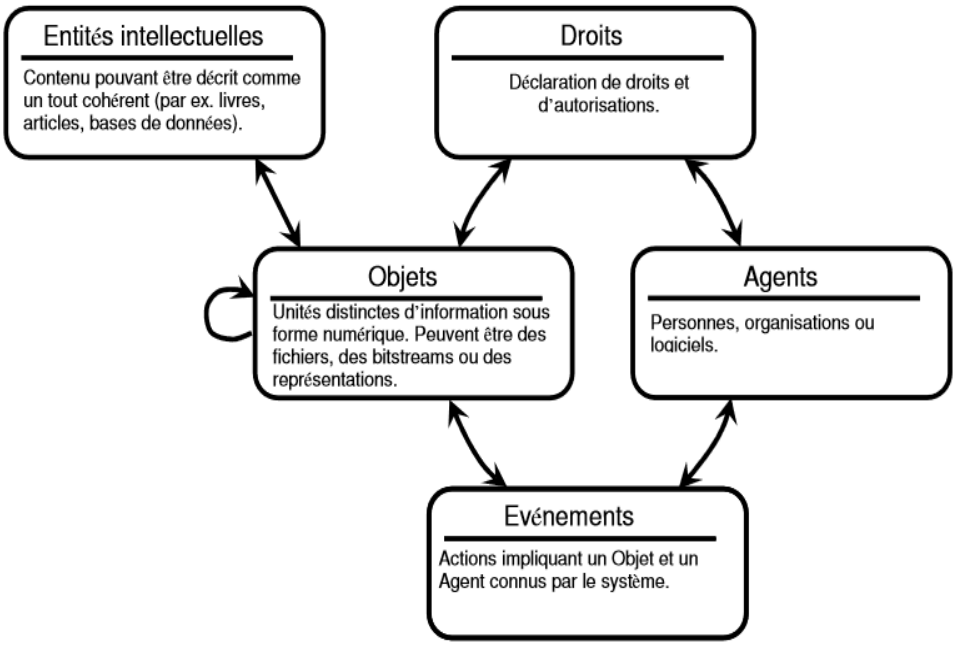
PREMIS définit des *unités sémantiques*, soit des éléments "d'information ou de connaissance" (Library of Congress Network Development and Marc Standards Office 2009, p. 7). Ces dernières peuvent être organisées sous forme de conteneurs regroupant des *unités sémantiques* "enfant". Par exemple, l'*unité sémantique objectIdentifier* regroupe deux *composants sémantiques* (*semantic components*) : *objectIdentifierType* et *objectIdentifierValue*. Ce sont ces deux unités qui seront renseignées et contenues dans l'unité supérieure. Une autre particularité liée aux *unités sémantiques* réside dans la possibilité d'utiliser des extensions. Soit des conteneurs permettant de référencer des métadonnées non PREMIS (Ibid., p. 8).

PREMIS définit au travers de son modèle de données cinq types d'éléments pouvant être mis en relation. Ces éléments, ou entités, regroupent les *unités sémantiques* correspondantes vues plus haut[28] (Harvey 2010, p. 74 ; Library of Congress Network Development and Marc Standards Office 2009, p. 9-12) :

- ***entités intellectuelles*** (*intellectual entities*) : ensembles de contenus considéré comme une unité intellectuelle distincte (ex : photographies, cartes). N'est pas défini dans PREMIS, car relève d'autres métadonnées (ex : EAD) ;
- ***objets*** (*objects*) : objets numériques préservés par le dépôt. PREMIS définit trois types d'objets, le fichier, la représentation et le train d'octet ;
- ***événements*** (*events*) : actions sur un objet dans l'entrepôt numérique (ex : migration, empreinte numérique) ;
- ***agents*** (*agents*) : personnes, organisations ou logiciels associés à un ou plusieurs évènements ;
- ***droits*** (*rights*) : déclarations relatives à une permission ou à un droit touchant un objet et restreignant éventuellement son usage.

[28] Ainsi *objectIdentifier* est une *unité sémantique* liée à l'entité *objets*.

Les entités du modèle PREMIS



(Library of Congress Network Development and Marc Standards Office 2009)

4.6 Qualités identifiées

Les qualités identifiées dans les textes normatifs évoqués plus haut ont été organisées selon un cadre conceptuel reprenant principalement les dimensions, sous-dimensions et indicateurs du projet QADEPs cité au point 3.2.4. En effet, ce projet étant spécifiquement lié aux archives électroniques, ce choix semblait le plus naturel. Malgré tout, certains éléments et arrangements ont été modifiés afin de correspondre au mieux aux définitions des dimensions de la qualité identifiées dans les textes normatifs. De même, le cadre conceptuel reprend, dans ses niveaux inférieurs, les concepts de variables, questions de mesures et niveaux de qualité définis dans la thèse de Basma Makhlouf Shabou étudiée au point 3.2.3. Le texte du projet QADEPs ne définissant que les dimensions de la qualité des archives, et non pas les niveaux inférieurs, ces derniers ont été en partie explicités à l'aide de la thèse envisagée plus haut et des différents textes vus au préalables. Les cadres conceptuels du présent travail, et leurs variables associées, sont présentés, respectivement pour la dimension de la preuve crédible et la dimension de l'exploitabilité, en annexe 2 et 3.

L'accent de cette analyse a été mis particulièrement sur les dimensions de la qualité de la preuve crédible et de l'exploitabilité, soit sur *l'Information de représentation* et l'*Information de pérennisation* du modèle de référence OAIS. C'est donc principalement au travers des métadonnées administratives que se fera cette analyse. Le schéma METS ne définissant ni vocabulaire ni syntaxe de métadonnées, les indicateurs et les variables associées seront mis en exergue en grande partie au travers du standard de métadonnées PREMIS. Malgré tout, certains éléments structurels ou de référencement du standard METS participent en partie à la mise en évidence d'éléments concourant à la mesure des dimensions de la qualité des archives électroniques.

4.6.1 Dimension de la preuve crédible

La preuve crédible est la "capacité d'un document à gagner la confiance de son utilisateur comme étant la source privilégiée à l'appui des faits" (Makhlouf Shabou 2013, p. 115). Cette première dimension de la qualité des archives est présente dans l'ensemble des textes normatifs analysés au travers de ses deux sous-dimensions que sont l'authenticité et la fiabilité. De façon générale, METS propose deux éléments pouvant référencer ou contenir des métadonnées spécifiques participant en partie à cette dimension, soit les éléments *<sourceMD>* qui enregistre les données sur le document original analogique et *<digiprovMD>* qui lui consigne les informations sur les modifications qu'a pu subir l'objet numérique tout au long de son cycle de vie, et ainsi juger d'éventuelles altérations qui pourraient dénaturer sa capacité à représenter de façon précise l'original dont il est issu (Digital Library Federation 2010, p. 42-44 ; Library of Congress 2011).

4.6.1.1 Sous-dimension de l'authenticité

L'authenticité est définie comme la capacité – pour un document – de prouver qu'il est bien ce qu'il prétend être, qu'il a bien été produit ou reçu par la personne qui prétend l'avoir produit ou reçu, et ce au moment où il prétend l'avoir été (Organisation internationale de normalisation 2001a, p. 7 ; Organisation internationale de normalisation 2011a, p. 3). Cette capacité probatoire doit pouvoir être garantie au travers de la mise en place de "politiques et [de] procédures de contrôle de la création, de la réception, de la transmission, de la préservation et de l'élimination de ceux-ci, de manière à garantir que leurs créateurs sont autorisés et identifiés comme tels, et qu'ils sont ainsi protégés contre des ajouts, des suppressions, des altérations, une exploitation ou une dissimulation non autorisés" (Organisation internationale de normalisation 2011a, p. 3). La norme OAIS préconise que les détenteurs de documents d'archives les gèrent de telle sorte que leur *Contenu d'information* et leur authenticité soient préservés (Consultative Committee for Space Data Systems 2012,

p. 2-1). A cette fin, elle demande la mise en place de stratégies et de procédures documentées "permettant la diffusion d'une information, copie authentifiée de l'original ou permettant de remonter à l'original" (Comité consultatif pour les systèmes de données spatiales, p. 3-1). Afin d'atteindre ce but, la norme OAIS s'intéresse tout particulièrement à deux indicateurs, soit l'identité et l'intégrité. Ces indicateurs correspondent en fait à l'*Information d'identification* et l'*Information d'intégrité* du modèle OAIS.

Indicateur de l'identité

L'identité est l'ensemble des caractéristiques qui permettent d'identifier un document ou document d'archives de façon unique, et de le distinguer des autres (InterPARES 2 Project 2012).

Dans sa définition du paquet d'information, la norme OAIS requière l'existence d'une *Information d'identification*[29] fournissant "un ou plusieurs identificateurs, ou systèmes d'identificateurs, grâce auxquels le Contenu d'information peut être identifié de façon unique" (Comité consultatif pour les systèmes de données spatiales 2005, p. 2-7). Dans le même ordre d'idée, le standard PREMIS propose et définit l'*unité sémantique objectIdentifier* comme "a designation used to uniquely identify the object within the preservation repository system in which it is stored" (PREMIS Editorial Committee 2012, p. 28). Afin de garantir cette identification unique, le standard recommande l'enregistrement d'identificateurs créés automatiquement par l'entrepôt numérique. La première variable liée à l'identité est la **Variable de l'identificateur**.

L'historique des identificateurs doit de même pouvoir être conservé. PREMIS indique dans son dictionnaire de métadonnées de préservation "a depositor might need to request a file by its original name. Also, the repository may need to reconstruct internal links for dissemination" (PREMIS Editorial Committee 2012, p. 74). Le dépôt numérique peut donc renommer l'objet numérique qui lui a été soumis, mais doit conserver son intitulé d'origine (Ibid). Une seconde variable liée à l'identité est la **Variable de l'identificateur original**.

Enfin, un dernier élément lié à l'indicateur d'identité est présenté sous l'angle du producteur. De façon générale, la norme OAIS demande la présence d'une *Information de provenance* décrivant "l'origine du Contenu d'information, qui en a eu la charge depuis sa création, ainsi que son historique" (Comité consultatif pour les systèmes de données spatiales 2005, p. 2-6). Si la connaissance du nom du producteur et celle des

[29] Cette *Information d'identification* est une partie de l'*Information de pérennisation*.

responsables de la conservation sont entre autre renseignées au travers des métadonnées descriptives (ISAAR pour le premier et ISDIAH pour le second) et de leurs équivalant en format d'encodage, les métadonnées administratives peuvent aussi avoir leur rôle à jouer dans cet aspect. En effet, le standard PREMIS s'intéresse lui plus précisément au mécanisme permettant de vérifier l'authenticité de l'objet numérique. A cette fin, PREMIS préconise, pour certains cas, l'utilisation et la conservation des informations de l'*unité sémantique signatureInformation* : "For archival storage, […] where digital signatures are used by the repository as a tool to confirm the authenticity of its stored digital objects over time, must the signature itself and the information needed to validate the signature be preserved" (PREMIS Editorial Committee 2012, p. 249). La dernière variable liée à l'identité est la **Variable de la signature**. Le signataire peut être identifié au travers de l'*unité sémantique signer*, ou à l'aide d'*agentIdentifier* s'il est un agent connu du dépôt numérique (Ibid., p. 103).

Indicateur de l'intégrité

Second indicateur de l'authenticité, "l'intégrité d'un document renvoie au caractère complet et non modifié de son état" (Organisation internationale de normalisation 2001a, p. 8). Afin de garantir cette intégrité, la mise en place de politiques et de procédures de gestion doit permettre de connaître quelles modifications peuvent être apportées à un document après sa création, par qui, et en quelles circonstances. En outre, ces modifications doivent pouvoir être documentées et traçables (Organisation internationale de normalisation 2011a, p. 3). Ce dernier point se rapporte plus précisément à la sous-dimension de la fiabilité et à son indicateur de la traçabilité.

La norme OAIS, dans sa définition du paquet d'information, demande la présence d'une *Information d'intégrité*[30], fournissant des outils de contrôle, afin de s'assurer que les objets numériques ne soient pas modifiés de façon non documentée (Consultative Committee for Space Data Systems 2012, p. 4-30). Elle précise encore "l'intégrité fournit un mécanisme ou un dispositif protecteur pour prémunir le Contenu d'information contre toutes altérations non documentées. Par exemple, il peut s'agir d'un checksum sur le Contenu d'information d'un Paquet d'informations numérique" (Comité consultatif pour les systèmes de données spatiales 2005, p. 2-7). PREMIS, de son côté, indique l'*unité sémantique objectCharacteristics* et plus particulièrement son composant *fixity* comme outil de vérification de l'intégrité. Des algorithmes de hachages permettent d'obtenir une empreinte numérique unique pour un objet

[30] Tout comme l'*Information d'identification*, cette *Information d'intégrité* est une partie de l'*Information de pérennisation.*

numérique donné. Si cette empreinte numérique est toujours identique par la suite, l'objet numérique n'a pas subi de modification dans le laps de temps donné. (PREMIS Editorial Committee 2012, p. 248). Afin de garantir cette intégrité, PREMIS préconise que soit effectué deux empreintes numériques utilisant deux algorithmes différents (Ibid). Il s'agit ici de la **Variable de l'empreinte numérique**.

La norme OAIS, en parlant d'intégrité, précise : "bien que l'Information d'intégrité au sein de l'Information de pérennisation d'un AIP garantisse que les bits associés à un Contenu d'information n'ont pas été altérés, il arrive un moment où ces bits ne sont plus sous une forme qui convient aux Utilisateurs" (Comité consultatif pour les systèmes de données spatiales 2005, p. 3-3). Dans ce cas l'*Archive* doit pouvoir migrer le *Contenu d'information* vers une nouvelle forme de représentation. Il s'agit ici d'une migration de formats. Lorsque cette circonstance se présente, l'OAIS devrait conserver les AIP originaux (Consultative Committee for Space Data Systems 2012, p. 3-3). Ce cas de figure, même s'il semble proche de l'indicateur d'intégrité, est à rapprocher de la sous-dimension de la fiabilité et de son indicateur de traçabilité. En effet, comme le stipule la norme ISO 15489, la traçabilité est le fait des opérations de création, d'enregistrement et de sauvegarde des données sur les mouvements et l'utilisation des documents (Organisation internationale de normalisation 2001a, p. 4). Même si l'opération de migration a un impact sur l'intégrité, la migration, dans notre propos, est à rapprocher de la traçabilité. Il s'agit en fait de ce que le dictionnaire PREMIS appelle un *événement*.

4.6.1.2 Sous-dimension de la fiabilité

La fiabilité – d'un document - est définie comme la capacité de son contenu à être considéré comme "la représentation complète et exacte des opérations, des activités ou des faits qu'il atteste, et sur lequel on peut s'appuyer lors d'opérations, d'activités ou de faits ultérieurs" (Organisation internationale de normalisation 2001a, p. 8). Deux indicateurs, dans le cadre conceptuel choisi, composent cette sous-dimension : la traçabilité et la complétude.

Indicateur de la traçabilité

La traçabilité est définie selon la norme ISO 15489 comme le "fait de créer, d'enregistrer et de préserver les données relatives aux mouvements et à l'utilisation des documents" (Organisation internationale de normalisation 2001a, p. 4). Comme précisé plus haut dans la section dévolue à la sous-dimension de l'authenticité, la norme OAIS préconise la mise en place de stratégies et de procédures documentées afin de garantir "la diffusion d'une information, copie authentifiée de l'original ou

permettant de remonter à l'original" (Comité consultatif pour les systèmes de données spatiales 2005, p. 3-1). Cette notion, liée à la traçabilité, est encore renforcée dans le cadre de l'*Information de provenance*, partie de l'*Information de pérennisation*. L'*Information de provenance* informe sur l'historique des *Contenus d'information*. Elle donne des informations sur son origine, les changements intervenus, et les différents acteurs qui en ont eu la garde. Cette information est garante, pour les utilisateurs, de la fiabilité des *Contenus d'information* (Consultative Committee for Space Data Systems 2012, p. 4-30). Cette notion de fiabilité est aussi reprise, de façon implicite et générale, par le standard METS. A travers son tutoriel publié par la Library of Congress, il est dit en parlant des métadonnées techniques, qu'elles sont une garantie de précision en ce qui concerne la fidélité d'une version numérisée face à son original (Library of Congress 2011, p. 1). C'est notamment l'élément *<sourceMD>* qui va renseigner ce fait, alors que l'élément *<digiprovMD>* s'intéressera lui, plus précisément, aux transformations liées à un objet électronique. Deux variables peuvent être comprises sous cet indicateur de la traçabilité : la variable du versement, qui renseigne sur la provenance, et la variable de l'enregistrement des modifications, qui renseigne sur les interventions qu'a subies l'objet numérique.

Le versement est un mode particulier de transmission. Il s'agit de la transmission d'AIP du *Producteur* à l'entité *Entrée*. La transmission, ou acquisition, peut être considérée comme le mouvement d'un document d'activité à travers l'espace ou le temps (InterPARES 2 Project 2012). Il s'agit donc ici de documenter le mouvement de l'objet numérique entre ces deux entités. Soit de mentionner l'origine et la date du versement et l'historique de conservation antérieur. Bien que la norme ISAD(G) permette de renseigner cette variable à l'aide des éléments d'information *Modalités d'entrée* et *Historique de la conservation*, et de prouver "son authenticité et son intégrité" à l'unité de description (Conseil international des archives 2000, p. 22-24), PREMIS propose de son côté l'*unité sémantique eventType* pour renseigner la **Variable du versement**. Ainsi, un vocabulaire spécifique peut informer sur l'acquisition d'un objet numérique par le dépôt. Ce vocabulaire peut être explicite ou formulé sous forme de codes. *Ingestion* peut de la sorte renseigner sur "[the] Process of adding objects to a Preservation Repository's storage system", alors que *Capture* peut être utilisé pour indiquer "a process by wich a Preservation Repository actively obtains Digital Objects for long term retention" (PREMIS Editorial Committee 2012, p. 134, 257, 260). Chaque institution

doit définir son propre vocabulaire. De nombreuses institutions, à l'instar des Archives d'État du Valais[31], utilisent *ingestion* pour renseigner la Variable du versement.

L'*Information de provenance* du modèle OAIS est responsable de l'enregistrement des traitements subis par l'objet numérique (Consultative Committee for space Data System 2012, p. 2-6, 4-30). Elle donne ainsi une certaine garantie quant à la fiabilité du *Contenu d'information*. Selon le dictionnaire PREMIS, "Actions that modify objects should always be recorded" (PREMIS Editorial Committee 2012, p. 130). Ces actions peuvent être de différentes natures : compression, décompression, migration, ou encore suppression d'un objet numérique (Ibid., p. 134, 135). Si certaines actions n'ont que peu d'impact sur l'objet, d'autres ont des conséquences importantes. La **Variable de l'enregistrement des modifications** doit donc être renseignée au travers de l'entité *événements*. Tout comme pour la variable du versement, l'*unité sémantique eventType* renseigne sur le type d'événements, dans ce cas précis de modifications, que subit l'objet numérique. De même, il s'agit d'utiliser un vocabulaire contrôlé ou un système de codes.

Indicateur de la complétude

La complétude est définie par le projet InterPARES comme "the characteristic of a record that refers to the presence within it of all the elements required by the creator and the juridical system for it to be capable of generating consequences" (InterPARES 2 Project 2012). Avec la naturalité et la primitivité, il s'agit d'une qualité requise pour constituer un document d'activité original (InterPARES 2 Project 2012 ; Makhlouf Shabou 2010, p. 72). Cette complétude sous-entend la capacité, pour le contenu d'un dossier, à renseigner de façon autonome sur une affaire spécifique (Makhlouf Shabou 2010, p. 158). La norme OAIS préconise, au travers d'un *Protocole de versement*, de vérifier la complétude des *Paquets d'information à verser* (Consultative Committee for Space Data Systems 2012, p. 2-9, 2-10). Cette complétude ne concerne pas que le *Contenu d'information*, mais bien tous les éléments nécessaires à la lecture et à la compréhension de ce dernier, soit les métadonnées associées. C'est notamment au travers des métadonnées que le contenu d'un dossier peut renseigner de façon autonome sur l'affaire dont il est rattaché. Ainsi, deux variables peuvent découler de cet indicateur de complétude : la **Variable complétude du Contenu d'information** et la **Variable complétude des métadonnées**. Seule section obligatoire de METS, la *carte de structure* permet d'organiser les différents contenus numériques présents en une structure cohérente, et faciliter par là leur compréhension (Digital Library

[31] Selon Alain Dubois, 20 février 2015.

Federation 2010, p. 57). Cependant, cet élément n'est pas suffisant pour permettre la mesure effective de la complétude. Cette dernière pourrait être mesurée, dans sa Variable complétude du *Contenu d'information* et sa Variable complétude des métadonnées, au travers d'un *Protocole de versement*. Les normes PAIMAS (*Producer Archive Interface Methodology Abstract Standard*) et PAIS (*Producer-Archive Interface Specification*) fournissent une méthodologie de travail et une représentation formelle des objets versés à l'*Archive* au travers des SIP (Banat-Berger, Duplouy, Huc 2009, p. 57-63). Cette représentation formelle, appelée aussi *Modèle des objets à transférer* permet de référencer les objets versés et de définir leurs caractéristiques et leurs relations. Grâce à une représentation dérivée d'un schéma XML, cet outil peut permettre une automatisation de la mesure de la complétude (Centre National d'Etudes Spatiales 2010 ; Banat-Berger, Duplouy, p. 57-63).

Malgré tout, cette variable est difficile à mesurer. La complétude doit effectivement s'intéresser aux processus de création des documents d'activité en amont de la chaîne de production, soit aux outils de Records Management et aux applications métiers[32].

4.6.1.3 Sous-dimension de la preuve historique

Dernière sous-dimension de la preuve crédible, la preuve historique repose sur trois indicateurs : la conformité légale, réglementaire et administrative, la rareté du témoignage, et enfin, l'étendue du témoignage. Cette sous-dimension n'étant pas définie explicitement, le recoupement de ces indicateurs et du cadre conceptuel de l'*Étude sur la définition et la mesure des qualités des archives définitives issues d'une évaluation*, permet de palier à ce défaut. Ainsi la preuve historique peut être définie comme étant la capacité d'un document à apporter un témoignage en regard de son contexte de création, de son exhaustivité, et de sa singularité. En effet, la conformité légale, réglementaire et administrative de notre cadre conceptuel se retrouve sous l'indicateur de niveau 2 "contexte de production" du texte cité ci-dessus qui englobe les indicateurs spécifiques "procédures de création" et "créateur". De son côté, la rareté du témoignage correspond à l'indicateur "rareté informationnelle" qui s'intéresse plus spécifiquement au contenu. Enfin, l'étendue du témoignage recoupe l'indicateur de niveau 1 "exhaustivité du témoignage", soit l'exhaustivité temporelle et thématique.

Si deux indicateurs (rareté du témoignage et étendue du témoignage) semblent relever d'une certaine subjectivité, ou du moins être difficilement mesurable, un autre offre un potentiel en termes de mesure : l'indicateur de la conformité légale, réglementaire et administrative.

[32] Selon Alain Dubois, 20 février 2015.

Indicateur de la conformité légale, réglementaire et administrative

Comme vu plus haut, la conformité légale, règlementaire et administrative correspond au "contexte de production" explicité dans la thèse intitulée *Étude sur la définition et la mesure des qualités des archives définitives issues d'une évaluation*. Ce contexte de production repose sur les procédures de création et sur le producteur (Makhlouf Shabou 2010, 118). Soit sur la personne physique ou morale qui crée, reçoit, ou assemble des documents d'activité dans le cadre de ses missions, fonctions ou activités, et sur les procédures qui amènent la création de documents d'activité (InterPARES 2 Project 2012). Les textes normatifs analysés dans ce travail ne donnent pas d'éléments concrets concernant cet indicateur. La **Variable conformité légale, réglementaire et administrative du producteur** et la **Variable conformité légale, réglementaire et administrative des procédures de création** sont à rechercher dans d'autres normes. Ainsi, la norme ISAAR permet, au travers des éléments d'information *Statut juridique* et *Fonctions et activités*, de renseigner la première variable liée au producteur (Conseil international des archives 2004, p. 17-19). Pour la seconde variable, liée aux procédures de création, ISDF propose l'élément d'information *Législation* qui identifie "la base légale de la fonction" associée (Conseil international des archives 2007, p. 18-20).

4.6.2 Dimensions de l'exploitabilité

L'exploitabilité est la capacité pour un document de pouvoir être localisé, récupéré, transmis et compris (Organisation internationale de normalisation 2001a, p. 8 ; Organisation internationale de normalisation 2011a, p. 3). La norme OAIS s'intéresse à cette dimension de la qualité à travers certaines de ses responsabilités. Ainsi l'*Archive* doit s'assurer de la compréhensibilité des documents d'archives, de leur pérennisation, et de leur statut juridique (Consultative Committee for Space Data systems 2012, p. 2-1, 3-1, 3-2). Soit des garanties couvrant les trois sous-dimensions de l'exploitabilité. METS préconise plusieurs éléments pouvant participer à la mesure de cette dimension de l'exploitabilité, *<sourceMD>* qui renseigne sur le document analogique qui serait la source du document numérique, *<techMD>* qui contribue à informer sur les spécifications techniques, et *<rightsMD>* qui indique les restrictions en termes de droits qui peuvent toucher l'objet numérique (Digital Library Federation 2010, p. 22). Cette dimension de la qualité des archives est formée de trois sous-dimensions : l'accessibilité technique, l'accessibilité cognitive, et l'accessibilité juridique.

4.6.2.1 Sous-dimension de l'accessibilité technique

Si l'accessibilité matérielle, dans un environnement analogique, correspond à la "facilité d'accéder au support des archives" (Makhlouf Shabou 2010, p. 122), l'accessibilité technique correspond, elle, dans un environnement électronique, à la facilité d'accéder au *Contenu d'information*. Il s'agit d'un point important dans cette analyse, car il est au centre de la problématique de l'archivage électronique. Une des responsabilités obligatoires de *l'Archive* découle justement de cette accessibilité. L'*Archive* doit "acquérir une maîtrise suffisante de l'information fournie […] pour pouvoir en garantir la Pérennisation" (Comité consultatif pour les systèmes de données spatiales 2005, p. 3-1). Pour remplir cette responsabilité, la norme OAIS requiert la mise en place de stratégies et procédures afin de garantir la conservation de l'information, et de se protéger de tout accident (Ibid.). Le cadre conceptuel du projet QADEPs propose trois indicateurs, soit la fonctionnalité des accès, la lisibilité, et la pérennisation des clés d'accès. Pour le présent travail, une organisation différente s'est imposée. Ainsi, les indicateurs retenus sont l'indicateur de la fonctionnalité des accès, l'indicateur de la lisibilité, et l'indicateur de la fonctionnalité et du rendu.

Indicateur de la fonctionnalité des accès

Premier indicateur de la sous-dimension de l'accessibilité technique, la fonctionnalité des accès est à rapprocher de la "repérabilité" vue dans l'*Étude sur la définition et la mesure des qualités des archives définitives issues d'une évaluation* (Makhlouf Shabou 2010, p. 122). La fonctionnalité des accès peut ainsi être définie comme l'habileté à repérer un document sur un support électronique. A cette fin, un système de localisation doit permettre d'accéder à l'objet électronique. La section de fichiers d'un document METS fournit un inventaire et la localisation du contenu numérique de l'objet décrit (Digital Library Federation 2010, p. 46). A travers l'élément *file location* <*FLocat*>, il donne une information de localisation (Ibid., p. 49). La **Variable du système de localisation** permet donc de mesurer cet indicateur de la fonctionnalité des accès.

Indicateur de la lisibilité

La lisibilité peut être traduite comme la capacité à lire un contenu du fait de son état physique et de la disponibilité d'un dispositif adéquat de lecture (Makhlouf Shabou 2010, p. 122). Cet indicateur revêt une importance primordiale dans un environnement électronique. En effet, si la principale préoccupation dans le monde analogique était de s'assurer que les supports d'information soient stables (Consultative Committee for Space Data Systems 2012, p. 2-1), la complexité des données électroniques, et la

multiplication des moyens d'y accéder demandent de s'assurer de la pérennité d'un certain nombre d'éléments indispensables à la restitution de l'information. Il y a donc lieu de s'assurer que la documentation nécessaire à la compréhension de l'environnement de l'objet numérique soit enregistrée. "Separating digital content from its environmental context can make the content unusable" (PREMIS Editorial Committee 2012, p. 245). Une représentation simplifiée de la répartition de l'information en couche peut être un point de départ pour identifier les variables pouvant être incluses dans cet indicateur. Pour rappel, le modèle présenté dans la section dévolue à l'archivage électronique à long terme, était composé de cinq couches : couche physique, couche binaire, couche codage, couche élément, et enfin couche objet (Banat-Berger, Duplouy, Huc 2009, p. 28-32).

La première couche, la couche physique correspond au support physique pour enregistrer les bits (Banat-Berger, Duplouy, Huc 2009, p. 28). La connaissance de cette information est importante pour choisir un support d'archivage adéquat et planifier ainsi les migrations (Ibid., p. 82, 120, 121). "Knowing the storage medium is an internal requirement in order to trigger preservation actions" (PREMIS Editorial Committee 2012, p. 79). PREMIS dispose d'une *unité sémantique* spécifique pour indiquer le support physique utilisé. Cette unité permet de savoir comment et quand procéder à la migration (Ibid.). Une première variable est ainsi liée à l'indicateur lisibilité, la **Variable medium de stockage**.

Dans le modèle OAIS, l'*Information de représentation* est l'information permettant d'expliciter un objet numérique (Consultative Committee for Space Data Systems 2012, p. 1-14). Cette dernière information, et plus particulièrement sa composante *Information de structure* correspond, de façon simplifiée, aux couches supérieures du modèle vue plus haut. Elle s'intéresse tout particulièrement aux "types de codage et […] modes d'organisation des différents granules d'information au sein d'un fichier numérique" (Banat-Berger, Duplouy, Huc 2009, p. 47), soit les formats de représentation, ou formats de données. Maintenir une documentation sur les formats permet d'appliquer les stratégies adéquates en termes de pérennisation. Il s'agit donc de disposer de métadonnées sur le format de fichier original, mais aussi sur l'environnement logiciel et matériel (Library of Congress Network Development and Marc Standards Office 2009, p. 3). PREMIS offre différentes unités sémantiques à même de renseigner ces éléments. Ainsi, la **Variable format de données** peut être mesurée à l'aide de l'*unité sémantique format* et de ses *composants sémantiques* (PREMIS Editorial Committee 2012, p. 53-62). L'identification du format peut notamment se faire à l'aide d'un nom ou d'un registre de format. Cette dernière façon

de faire permet l'enregistrement d'une information détaillée (Ibid., p. 53). Pour plus de sécurité, deux registres de formats devraient être utilisés. En effet, aucune certitude n'est pour l'heure établie quant à la pérennité et la stabilité de ces registres de format[33].

Une autre variable fortement liée au format de données, la **Variable application de création** permet de renseigner sur l'application ayant servi à la création de l'objet numérique. Deux raisons peuvent être mises en évidence pour s'intéresser à cette variable. Premièrement, l'application de création peut, dans certains cas, être utilisée comme application de restitution. Deuxièmement, certaines versions de logiciels sont connues pour créer des erreurs de conversions (PREMIS Editorial Committe 2012, p. 63). Le renseignement de l'*unité sémantique creatingApplication* et ses *composants sémantiques* est donc un gage de pérennisation.

Une des responsabilités de l'*Archive* selon le modèle OAIS est de garantir la pérennisation de l'information conservée (Roussel 2012-2014, p. 62, 63). Cette responsabilité ne peut se faire qu'en documentant les moyens de restitution de cette dernière. Connaître l'environnement logiciel, voire matériel, capable de restituer et d'exploiter l'information numérique est donc central en termes de pérennisation de l'information (PREMIS Editorial Committee 2012, p. 245). PREMIS propose l'*unité sémantique environment* et ses *composants sémantiques* afin de disposer de cette information. La **Variable environnement de restitution** est ainsi une composante essentielle dans la mesure de l'accessibilité technique. A l'exemple du format, l'environnement de restitution devrait être documenté au travers d'un registre (Ibid., p. 80).

Indicateur de la fonctionnalité et du rendu

Les stratégies de pérennisation, comme la migration de formats, peuvent avoir des effets néfastes en termes de conservation numérique. En plus de poser certains problèmes au niveau de l'intégrité et de l'authenticité, des dangers de pertes de fonctionnalités risquent d'apparaître dans l'application de stratégies de pérennisation (Deljanin 2012, p. 51). Les caractéristiques significatives peuvent être définies de la façon suivante : "the characteristics of digital objects that must be preserved over time in order to ensure the continued accessibility, usability, and meaning of the objects, and their capacity to be accepted as evidence of what they purport to record" (Wilson 2008, p. 15). Si cet indicateur est aussi en partie un composant de l'authenticité, ses

[33] Une modification de la valeur du champ *Unique ID* du registre PRONOM aurait même été constatée. Entretien avec Emmanuel Ducry des Archives d'État de Genève, 24 février 2015.

caractéristiques concourant à garantir l'accessibilité et l'utilisabilité le rapproche de la dimension de l'exploitabilité et de l'indicateur de la fonctionnalité et du rendu. PREMIS offre la possibilité d'indiquer les caractéristiques d'un objet numérique que l'*Archive* considère comme suffisamment importantes pour être préservées (PREMIS Editorial Committee 2012, p. 39). La **Variable caractéristiques significatives** peut donc être mesurée par l'*unité sémantique significantProperties*.

4.6.2.2 Sous-dimension de l'accessibilité cognitive

Deuxième sous-dimension de l'exploitabilité, l'accessibilité cognitive correspond à la "facilité d'accéder au contenu des archives et à leur contexte de création" (Makhlouf Shabou 2010, p. 120). Deux indicateurs se situent sous cette sous-dimension : la repérabilité et la compréhensibilité.

Indicateur de la repérabilité

Selon l'*Étude sur la définition et la mesure des qualités des archives définitives issues d'une évaluation*, la repérabilité, ou identifiabilité, peut être comprise comme la capacité à repérer un objet archivé (Makhlouf Shabou p. 121). Cette définition précise encore qu'elle repose sur "la structure classificatoire logique qui sert à organiser intellectuellement les documents, tel le plan de classification, et d'autre part le titre attribué au dossier d'archives" (Schellenberg 1965, cité dans Makhlouf Shabou 2010, p. 121 ; Makhlouf Shabou 2010, p. 121). La norme OAIS propose une *Information d'identification* devant permettre l'identification du *Contenu d'information* (Consultative Committee for Space Data Systems 2012, p. 1-14). La mesure de cette repérabilité, dans son acceptation cognitive, peut être effectuée à l'aide d'éléments présents dans les normes du Conseil international des archives. Ainsi la **Variable de l'identifiant intellectuel**, liée à la structure classificatoire, et la **Variable de l'intitulé** peuvent être mesurées à l'aide des éléments d'information *Référence* et *Intitulé* de la norme ISAD(G) (Conseil international des archives 2000, p. 14, 15).

Indicateur de la compréhensibilité

La compréhensibilité est la capacité et l'aisance à appréhender le *Contenu d'information*. (Makhlouf Shabou 2010, p. 121). La norme OAIS cite, comme responsabilité obligatoire, de devoir s'assurer que l'information est immédiatement compréhensible par la *Communauté d'utilisateurs cible*, et ce sans assistance (Consultative Committee for Space Data Systems 2012, p. 3-1). OAIS demande à cette fin l'implication d'une *Information de contexte* et d'une *Information de provenance* en complément du *Contenu d'information*. L'*Information de contexte* "documente les

relations entre le Contenu d'information et son environnement, y compris les motifs de création du Contenu d'information et les relations qu'il entretient avec d'autres Objets-contenu d'information existant par ailleurs", alors que l'*Information de provenance* s'intéresse à l'historique des *Contenus d'information*. Dans ce sens, elle apporte un témoignage sur l'origine, les modifications intervenues, et les responsables de sa conservation (Ibid., p. 4-30). Tout comme pour l'indicateur de la traçabilité, la compréhensibilité est un indicateur qui peut être mesuré, pour une grande partie, à l'aide des normes du Conseil international des archives. Ainsi, la **Variable contexte de création**, dans sa facette de la compréhensibilité, peut être mesurée à l'aide de la norme ISDF et de son élément d'information *Législation*. La législation est en effet à l'origine de la fonction responsable de la création d'un document d'archive (Conseil international des archives 2007, p. 18). La **Variable origine du Contenu d'information**, quant à elle, peut être renseignée sur la base de l'élément d'information *Nom du producteur* de la norme ISAD(G). Cet élément peut être enregistré dans un fichier d'autorité (Conseil international des archives 2000, p. 18). L'élément d'information *Historique de la conservation*, présent lui aussi dans ISAD(G), peut de son côté servir comme moyen de pondération pour la **Variable historique de la conservation** (Ibid., p. 22). Ce dernier élément renseigne sur les différents changements de dépositaires qui ont pu avoir lieu.

4.6.2.3 Sous-dimension de l'accessibilité juridique

Dernière sous-dimension proposée dans cette analyse, l'accessibilité juridique peut être entendue comme la "communicabilité des archives qui s'appuie sur l'autorisation juridique de consulter les documents d'archives" (Makhlouf Shabou 2010, p. 121). Cette sous-dimension se base ainsi sur l'indicateur de la communicabilité.

Indicateur de la communicabilité

Dans sa thèse, Basma Makhlouf Shabou définit la communicabilité comme la possibilité de consultation basée sur des textes légaux ou réglementaire (Makhlouf Shabou 2010, p. 121). En reprenant la définition générale de l'exploitabilité, le terme "utilisation des archives" est mis en évidence (Makhlouf Shabou 2010, p. 120). L'accessibilité juridique doit donc inclure cette notion d'utilisation dans sa définition. En effet, certains documents pouvant être communiqués peuvent détenir des modalités limitant leur usage ou leur utilisation. La norme OAIS indique que les droits de propriétés intellectuels doivent être transférés à l'*Archive*, ou si ce n'est pas le cas, que ces derniers doivent être spécifiés de façon à définir d'éventuelles restrictions découlant des ayants droits (Consultative Committee for Space Data Systems 2012, p.

3-2). Certains objets numériques peuvent de même n'être diffusés qu'auprès de groupes d'utilisateurs autorisés, dans ce cas, *l'Archive* "doit rendre publique sa politique d'accès et les restrictions induites, afin de protéger les droits de chacune des parties" (Ibid., p. 3-6).

Plusieurs types de restrictions peuvent limiter la communicabilité des archives. Des délais de protection peuvent tout d'abord restreindre l'accès aux archives. Ainsi, le Canton de Vaud, à l'instar des autres cantons, applique des délais de protection ordinaire de 30 ans et des délais de protection spéciaux pouvant courir jusqu'à cent ans après l'ouverture d'un dossier (Canton de Vaud. Grand Conseil 2011). Ensuite, des restrictions d'accès peuvent venir de lois protégeant le droit d'auteur (Canton de Vaud. Grand Conseil, 2002 ; Confédération suisse 2011).

Dans le premier cas, le texte de loi qui implique une restriction peut être renseigné au travers de l'*unité sémantique otherRightsInformation* et de ses *composants sémantiques*, alors que la durée du délai de protection sera renseignée sous l'*unité sémantique termOfRestriction* (PREMIS Editorial Committee 2012, p. 213-222). La **Variable protection des données** peut donc être mesurée au travers ces éléments.

Pour le second cas, la **Variable protection intellectuelle*,*** peut trouver dans l'*unité sémantique copyrightInformation* et ses *composants sémantiques* des éléments spécifiques de mesures.

4.7 Conclusion sur les qualités identifiées dans les normes et standards

Cette analyse a permis d'identifier différentes variables participant à la mesure des dimensions de la qualité des archives électroniques, et mises en avant par les normes et standards sélectionnées. Ainsi 23 variables ont été identifiées. Dix sont liées à la dimension de la preuve crédible, et 13 à la dimension de l'exploitabilité. Comme indiqué plus haut, la dimension de la représentativité a été écartée en raison de son caractère subjectif. Parmi ces 23 variables, sept seraient mesurables à l'aide d'éléments d'information des normes du Conseil international des archives (ISAD(G), ISAAR, ISDF et ISDIAH), et deux à l'aide des normes PAIMAS et PAIS.

La norme OAIS offre une vision conceptuelle de la problématique en s'affranchissant de tout contexte, application ou domaine spécifique (Banat-Berger, Duplouy, Huc 2009, p. 41). Dans ce sens, elle est une base pour définir ce que doit intégrer l'*Archive* afin de répondre aux niveaux de qualités adéquats. De son côté, METS, qui permet de référencer des fichiers de données et ses métadonnées associées (Ibid., p. 64), offre

des éléments structurels permettant de mesurer certains indicateurs de la qualité. PREMIS, quant à lui, propose des *unités sémantiques* traduisant le modèle OAIS et permettant la mesure de l'intégrité, de la lisibilité et de la traçabilité des interventions sur un objet numérique (PREMIS Editorial Committee 2012, p. 3 ; Library of Congress Network Development and Marc standards Office 2009, p. 3, 4). Ce dernier texte normatif est donc un élément clé pour l'opérationnalisation des variables identifiées.

5. Opérationnalisation

Sur la base des dimensions, sous-dimensions et indicateurs identifiés plus haut, le présent chapitre propose l'opérationnalisation, soit la création de variables, questions de mesures et niveaux de qualité. Les grilles de mesures ont été construites sur le modèle du cadre conceptuel du projet QADEPs et de l'*Étude sur la définition et la mesure des qualités des archives définitives issues d'une évaluation* vues au point 3.2.4 et 3.2.3. Ainsi, les trois premiers niveaux, soit la dimension, la sous-dimension, et l'indicateur permettent de situer la variable de façon hiérarchique. Une question et des niveaux de qualité ont été associés à cette dernière afin de permettre une mesure. Des notes d'application permettent de préciser les différents éléments vus plus haut.

Afin de juger de la capacité d'automatisation, une instance XML permet de situer l'élément de mesure et sa valeur.

1 Preuve crédible

1.1 Authenticité

1.1.1 Identité

Un identificateur permet-il de distinguer de façon unique l'objet numérique ?

Niveaux de qualité		**Variable de l'identificateur**
	1	Identificateur inexistant
	2	Identificateur existant et créé par un système extérieur au dépôt numérique
	3	Identificateur existant et créé automatiquement par le dépôt numérique

Exemple :

```
<premis:objectIdentifier>
        <premis:objectIdentifierType>hdl</premis:objectIdentifierType>
        <premis:objectIdentifierValue>am.dlib/.09601</premis:objectIdentifierValue></p
remis:objectIdentifier>
```

Dans la présente opérationnalisation, l'*unité sémantique objectIdentifier* et ses deux composants *objectIdentifierType* et *objectIdentifierValue* permettent la mesure de la variable.

Un identificateur se doit d'être unique. A cette fin, il doit être créé de façon automatique par le dépôt numérique. Un identificateur standardisé externe au dépôt, de type ISBN ou issu d'un cadre de classification, n'est pas suffisant (PREMIS Editorial Committee 2012, p. 28).

1 Preuve crédible

1.1 Authenticité

1.1.1 Identité

Un identificateur d'origine permet-il de distinguer l'objet numérique ?

Niveaux de qualité	**Variable de l'identificateur original**
1	Identificateur d'origine inexistant
2	Identificateur d'origine existant

Exemple :

<premis:originalName>ico0025.tif</premis:originalName>

Dans la présente opérationnalisation, l'*unité sémantique originalName* permet la mesure de la variable.

1 Preuve crédible

1.1 Authenticité

1.1.1 Identité

Une signature valide permet-elle d'associer l'objet numérique au producteur ?

Niveaux de qualité	**Variable de la signature**
1	Signature inexistante
2	Signature existante
3	Signature existante avec indication du nom de l'individu, de l'institution ou de l'autorité responsable

Exemple :

```
<premis:signatureInformation></premis:signatureInformation>
        <premis:signature>
                <premis:signatureEncoding>Base64</premis:signatureEncoding>
                <premis:signer>AMJT</premis:signer>
                <premis:signatureMethod>DSA-SHA1</premis:signatureMethod>
                <premis:signatureValue>juS5R884qoFR… </premis:signatureValue>
        <premis:signatureValidationRules>…</premis:signatureValidationRules>
                <premis:signatureProperties>…</premis:signatureProperties
                <premis:keyInformation>…</premis:keyInformation>
        </premis:signature>
</premis:signatureInformation>
```

Dans la présente opérationnalisation, l'*unité sémantique signatureInformation*, son composant sémantique *signature* avec leurs propres composants sémantiques

signatureEncoding, *signer*, *signatureMethod*, *signatureValue*, *signatureValidationRules*, *signatureProperties* et *keyInformation* permettent la mesure complète de la variable. *signer* permet de renseigner sur l'identité de l'individu, de l'institution ou de l'autorité responsable. Si ces derniers ont un rôle d'*agent*, l'*unité sémantique agentIdentifier* peut être utilisée.

Dans le cas de versement au dépôt numérique ou de diffusion du dépôt numérique, la validation se fait peu après la signature. La conservation de la signature n'est donc pas requise. Cette dernière est requise lorsque le dépôt numérique veut utiliser ces dernières comme outil de confirmation de l'authenticité des objets numériques à travers le temps nécessaire (PREMIS Editorial Committee 2012, p. 249).

---

1 Preuve crédible

1.1 Authenticité

1.1.2 Intégrité

Une fonction algorithmique permet-elle l'enregistrement d'une empreinte numérique ?

Niveaux de qualité	**Variable de l'empreinte numérique**
1	Empreinte numérique inexistante
2	Empreinte numérique existante
3	Deux empreintes numériques

Exemple :

```
<premis:fixity>
        <premis:messageDigestAlgorithm>MD5</premis:messageDigestAlgorithm>
        <premis:messageDigest>36b03197ad066cd719906c55eb68ab8d</premis:m
        essageDigest>
        <premis:messageDigestOriginator>LocalDCMS</premis:messageDigestOrigi
        nator>
</premis:fixity>
```

Dans la présente opérationnalisation, l'*unité sémantique fixity* et ses trois composants sémantiques *messageDigestAlgorithm*, *messageDigest* et *messageDigestOriginator* permettent la mesure de la variable.

La pratique recommandée veut que deux empreintes numériques soient calculées à partir de deux algorithmes différents (PREMIS Editorial Committee 2012, p. 47). A noter que l'acte de procéder à un contrôle d'intégrité est considéré par PREMIS comme un *événement*, et son résultat comme informations sur le résultat d'un *événement*. Par conséquent, seuls les composants *messageDigestAlgorithm* et *messageDigest* sont requis sous l'*unité sémantique objectCharacteristics* pour une comparaison future (Ibid.).

---

1 Preuve crédible

1.2 Fiabilité

1.2.1 Traçabilité

Une information permet-elle de renseigner le versement de l'objet numérique au dépôt numérique ?

Niveaux de qualité	**Variable du versement**
1	Information concernant le versement inexistante
2	Information concernant le versement renseignée

Exemple :

```
<premis:event>
    …
    <premis:eventType>ingestion</premis:eventType>
    <premis:eventDateTime>2015-02-06T00:00:00.002</premis:eventDateTime>
    <premis:eventDetail>ingester1_0.exe</premis:eventDetail>
    …
</premis:event>
```

Dans la présente opérationnalisation, *eventType* devrait être accompagnée de l'ensemble des unités sémantiques de l'entité *Event*, soit *eventIdentifier*, *eventDateTime*, *eventDetail*, *eventOutComeInformation*, *linkingAgentIdentifier* et *linkingObjectIdentifier*.

Un vocabulaire spécifique permet de renseigner cette variable. Ainsi *capture* et/ou *ingestion* permettent de renseigner le versement au dépôt numérique au travers de l'*unité sémantique eventType* (PREMIS Editorial Committee 2012, p. 134, 135).

1 Preuve crédible

1.2 Fiabilité

1.2.1 Traçabilité

Une information permet-elle de renseigner les modifications subies par l'objet numérique ?

Niveaux de qualité	**Variable de l'enregistrement des modifications**
1	Information concernant les modifications inexistante
2	Information concernant les modifications existante

Exemple :

```
<premis:event>
    …
    <premis:eventType>deletion</premis:eventType>
    <premis:eventDetail>object permanently delete by request of…
    </premis:eventDetail>
```

…
<premis:event>

Dans la présente opérationnalisation, *eventType* devrait être accompagnée de l'ensemble des unités sémantiques de l'entité *Event*, soit *eventIdentifier*, *eventDateTime*, *eventDetail*, *eventOutComeInformation*, *linkingAgentIdentifier* et *linkingObjectIdentifier*.

Il est recommandé de renseigner de façon générale le type d'événement dans *eventType* et d'utiliser l'*unité sémantique eventDetail* pour compléter celle-là (PREMIS Editorial Committee 2012, p. 135).

---

1 Preuve crédible

1.2 Fiabilité

1.2.2 Complétude

Une information permet-elle de renseigner sur la présence de l'ensemble des éléments requis du Contenu d'information ?

Variable complétude du *Contenu d'information*

Cette variable n'est opérationnalisable qu'à l'aide d'un *Protocole de versement* comprenant un *Modèle des objets à transférer*. La mesure pourra se faire en regard des objets réellement versés et de ce modèle (Centre National d'Etudes Spatiales 2010 ; Banat-Berger, Duplouy, p. 57-63).

---

1 Preuve crédible

1.2 Fiabilité

1.2.2 Complétude

Une information permet-elle de renseigner sur la présence de l'ensemble des métadonnées requises pour décrire le Contenu d'information ?

Variable complétude des métadonnées

Cette variable n'est opérationnalisable qu'à l'aide d'un *Protocole de versement* comprenant un *Modèle des objets à transférer*. La mesure pourra se faire en regard des objets réellement versés et de ce modèle (Centre National d'Etudes Spatiales 2010 ; Banat-Berger, Duplouy, p. 57-63).

---

1 Preuve crédible

1.3 Preuve historique

1.3.1 Conformité légale, réglementaire et administrative

Une information permet-elle de renseigner sur le statut juridique, les fonctions et les activités du producteur ?

Variable conformité légale, réglementaire et administrative du producteur

Cette variable n'est opérationnalisable qu'à l'aide de la norme ISDF. La mesure pourra se faire en regard des éléments d'information *Statut juridique* et *Fonctions et activités* (Conseil international des archives 2004, p. 17-19).

---

1 Preuve crédible

1.3 Preuve historique

1.3.1 Conformité légale, réglementaire et administrative

Une information permet-elle de renseigner sur la législation qui a amené la création, la réception ou l'assemblement du Contenu d'information?

Variable conformité légale, réglementaire et administrative des procédures de création

Cette variable n'est opérationnalisable qu'à l'aide de la norme ISDF. La mesure pourra se faire en regard de l'élément d'information *Législation* (Conseil international des archives 2007, p. 18-20).

---

2 Exploitabilité

2.1 Accessibilité technique

2.1.1 Fonctionnalité des accès

Une information permet-elle d'indiquer la localisation de l'objet numérique ?

Niveaux de qualité	**Variable du système de localisation**
1	Information concernant le système de localisation inexistante
2	Information concernant le système de localisation renseignée

Exemple :

```
<mets:fileSec>
        <mets:fileGrp USE="MASTER">
                        <mets:file MIMETYPE="image/tiff">
                <METS:FLocat LOCTYPE="URL"
        xlink:href="file:///2010_avm_0135_Glion/Projet_20100512.docx"/>
                        </mets:file>
        </mets:fileGrp>
</mets:fileSec>
```

Dans la présente opérationnalisation, l'élément *file location* (*FLocat*) de la *section de fichier* doit être renseigné.

2 Exploitabilité

2.1 Accessibilité technique

2.1.2 Lisibilité

Une information permet-elle de renseigner sur le support de stockage utilisé ?

Niveaux de qualité	**Variable medium de stockage**
1	Information concernant le support de stockage inexistante
2	Information concernant le support de stockage renseignée

Exemple :

```
<premis:storage>
        …
        <premis:storageMedium>disk</premis:storageMedium>
</premis:storage>
```

Dans la présente opérationnalisation, *storageMedium* doit être renseigné.

2 Exploitabilité

2.1 Accessibilité technique

2.1.2 Lisibilité

Une information permet-elle de renseigner sur le format de données de l'objet numérique ?

Niveaux de qualité	**Variable format de données**
1	Information concernant le format de données inexistante
2	Information concernant le format de données renseignée
3	Information concernant le format de données renseignée et référencée à l'aide d'un registre de formats extérieur
4	Information concernant le format de données renseignée et référencée à l'aide de plusieurs registres de formats extérieur

Exemple :

```
<premis:format>
        <premis:formatDesignation>
                <premis:formatName>image/tiff</premis:formatName>
```

```
            <premis:formatVersion>6.0</premis:formatVersion>
        </premis:formatDesignation>
        <premis:formatRegistry>
            <premis:formatRegistryName>PRONOM
            </premis:formatRegistryName>
            <premis:formatRegistryKey>fmt/10</premis:formatRegistryKey>
            <premis:formatRegistryRole>specification
            </premis:formatRegistryRole>
        </premis:formatRegistry>
</premis:format>
```

Dans le cas où le format de données n'est pas reconnu, il y a lieu d'indiquer ce dernier fait et de tenter une identification manuelle (PREMIS Editorial Committee 2012, p. 53). L'identification du format à l'aide d'un registre de format extérieur est une plus-value en termes de renseignements (Ibid., p. 58). Deux registres de formats de fichiers ou plus sont une plus-value en termes de pérennisation.

Diverses recommandations existent en ce qui concerne les formats candidats pour l'archivage électronique à long terme. Ainsi les critères d'ouverture, d'indépendance et de diffusion sont importants à prendre en compte. Malgré tout, chaque institution doit, sur la base d'une analyse des risques et de son environnement, définir un set de format dédié à l'archivage électronique à long terme. Dans ce sens, ces critères n'ont pas été intégrés à l'opérationnalisation de cette variable.

Dans la présente opérationnalisation, *format* devrait être accompagné de ses *composants sémantiques formatDesignation* et *formatRegistry*.

---

2 Exploitabilité

2.1 Accessibilité technique

2.1.2 Lisibilité

Une information permet-elle de renseigner sur l'application de création de l'objet numérique ?

Niveaux de qualité	**Variable application de création**
1	Information concernant l'application de création inexistante
2	Information concernant l'application de création renseignée
3	Information concernant l'application de création renseignée avec la date de création de l'objet numérique

Exemple :

```
<premis:creatingApplication>
        <premis:creatingApplicationName>Gimp</premis:creatingApplicationName>
        <premis:creatingApplicationVersion>2.8</premis:creatingApplicationVersion>
        <premis:dateCreatedByApplication>201502141030</premis:dateCreatedByA
        pplication>
```

```
</premis:creatingApplication>
```

Dans la présente opérationnalisation, *creatingApplication* devrait être accompagné de ses *composants sémantiques creatingApplicationName*, *creatingApplicationVersion*, et *dateCreatedByApplication*.

2 Exploitabilité

2.1 Accessibilité technique

2.1.2 Lisibilité

Une information permet-elle de renseigner sur l'environnement logiciel et/ou matériel permettant la restitution et l'exploitation de l'objet numérique ?

Niveaux de qualité	**Variable environnement de restitution**
1	Information concernant l'environnement inexistante
2	Information concernant l'environnement renseignée
3	Information concernant l'environnement renseignée et référencée à l'aide d'un registre extérieur

Exemple :

```
<premis:environment>
        <premis:environmentCharacteristic>recommended</premis:environmentChar
        acteristic>
        <premis:environmentPurpose>render</premis:environmentPurpose>
        <premis:environmentPurpose>edit</premis:environmentPurpose>
        <premis:software>
                <premis:swName>Adobe Acrobat</premis:swName>
                <premis:swVersion>5.0</premis:swVersion>
                <premis:swType>renderer</premis:swType>
        </premis:software>
        <premis:software>
                <premis:swName>Windows</premis:swName>
                <premis:swVersion>XP</premis:swVersion>
                <premis:swType>operatingSystem</premis:swType>
        </premis:software>
        <premis:hardware>
                <premis:hwName>Intel x86</premis:hwName>
                <premis:hwType>processor</premis:hwType>
        </premis:hardware>
</premis:environment>
```

L'identification de l'environnement de restitution à l'aide d'un registre extérieur similaire à celui des formats est une plus-value en termes de renseignements (PREMIS Editorial Committee 2012, p. 80).

Dans la présente opérationnalisation, *environment* devrait être accompagné de ses

composants sémantiques environmentPurpose, et *software*.

2 Exploitabilité

2.1 Accessibilité technique

2.1.3 Fonctionnalité et rendu

Une information permet-elle de renseigner sur les caractéristiques primordiales de l'objet numérique que se doit de préserver l'Archive ?

Niveaux de qualité	**Variable caractéristiques significatives**
1	Information concernant les caractéristiques significatives inexistante
2	Information concernant les caractéristiques significatives renseignée
3	Information concernant les caractéristiques significatives renseignée et testée après l'application des stratégies de préservation

Exemple :

```
<premis:significantProperties>
        <premis:significantPropertiesType> behavior
        </premis:significantPropertiesType>
        <premis:significantPropertiesValue> hyperlinks traversable
        </premis:significantPropertiesValue>
</premis:significantProperties>
```

significantProperties peut être utilisé comme mesure de l'efficacité d'une méthode de préservation telle que la migration. Si une caractéristique significative n'est pas maintenue après une migration, le processus de migration peut être considéré comme défaillant (PREMIS Editorial Committee 2012, p. 39).

Dans la présente opérationnalisation, *significantProperties* devrait être accompagné de ses *composants sémantiques significantPropertiesType*, et *significantPropertiesValue*.

2 Exploitabilité

2.2 Accessibilité cognitive

2.2.1 Repérabilité

Une identifiant permet-il de situer l'objet archivé dans un système classificatoire ?

Variable de l'identifiant intellectuelle

Cette variable n'est opérationnalisable qu'à l'aide de la norme ISAD(G). La mesure

pourra se faire en regard de l'élément d'information *Référence* (Conseil international des archives 2000, p. 14, 15).

2 Exploitabilité

2.2 Accessibilité cognitive

2.2.1 Repérabilité

Un intitulé permet-il de situer l'objet archivé dans un système classificatoire ?

Variable de l'intitulé

Cette variable n'est opérationnalisable qu'à l'aide de la norme ISAD(G). La mesure pourra se faire en regard de l'élément d'information *Intitulé* (Conseil international des archives 2000, p. 14, 15).

2 Exploitabilité

2.2 Accessibilité cognitive

2.2.2 Compréhensibilité

Une information permet-elle de comprendre le motif de la création de l'objet archivé ?

Variable contexte de création

Cette variable n'est opérationnalisable qu'à l'aide de la norme ISDF. La mesure pourra se faire en regard de l'élément d'information *Législation* (Conseil international des archives 2007, p. 18).

2 Exploitabilité

2.2 Accessibilité cognitive

2.2.2 Compréhensibilité

Une information permet-elle de situer l'origine de l'objet archivé ?

Variable origine du Contenu d'information

Cette variable n'est opérationnalisable qu'à l'aide de la norme ISAD(G). La mesure pourra se faire en regard de l'élément d'information *Nom du producteur* (Conseil international des archives 2000, p. 18).

2 Exploitabilité

2.2 Accessibilité cognitive

2.2.2 Compréhensibilité

Une information permet-elle de renseigner sur l'historique de la conservation de l'objet archivé ?

Variable historique de la conservation

Cette variable n'est opérationnalisable qu'à l'aide de la norme ISAD(G). La mesure pourra se faire en regard de l'élément d'information *Historique de la conservation* (Conseil international des archives 2000, p. 22).

---

2 Exploitabilité

2.3 Accessibilité juridique

2.3.1 Communicabilité

Des restrictions de protection des données sont-elles liées à l'objet numérique ?

Niveaux de qualité	**Variable protection des données**
1	Restrictions à durée indéterminée empêchant la diffusion de l'objet numérique
2	Restrictions à durée déterminée empêchant la diffusion de l'objet numérique
3	Aucune restriction sur l'objet numérique

Exemple :

```
<premis:rightsStatement>
        …
        <premis:rightsBasis>other</premis:rightsBasis>
                <premis:otherRightsInformation>
                …
                </premis:otherRightsInformation>
                <premis:rightsGranted>
                <act>Disseminate</act>
                <restriction>Disallow</restriction>
                <termOfRestriction>
                        <startDate>20140225</startDate>
                        <endDate>20440225</endDate>
                </termOfRestriction>
                <rightsGrantedNote>Soumis à un délai de protection de 30
                ans</rightsGrantedNote>
                </premis:rightsGranted>
        …
</premis:rightsStatement>
```

Dans la présente opérationnalisation, l'*unité sémantique otherRightsInformation* devrait être accompagnée des *unités sémantiques act* et *restriction,* alors que *l'unité sémantique termOfGrant* englobera *termOfRestriction*, *startDate*, *endDate.*

---

2 Exploitabilité
2.3 Accessibilité juridique
2.3.1 Communicabilité

Des restrictions de droit d'auteur sont-elles liées à l'objet numérique ?

Niveaux de qualité	**Variable protection intellectuelle**
1	Restrictions à durée indéterminée empêchant la diffusion et/ou l'utilisation de l'objet numérique
2	Restrictions à durée déterminée empêchant la diffusion et/ou l'utilisation de l'objet numérique
3	Aucune restriction sur l'objet numérique

Exemple :

```
<premis:rightsStatement>
        …
        <premis:rightsBasis>copyright</premis:rightsBasis>
                <premis:copyrightInformation>
                        <premis:copyrightStatus>copyrighted
                        </premis:copyrightStatus>
                        <premis:copyrightStatusDeterminationDate>20150809
                        </premis:copyrightStatusDeterminationDate>
                </premis:copyrightInformation>
</premis:rightsStatement>
```

Dans la présente opérationnalisation, *copyrightInformation* devrait être accompagnée de ses *composants sémantiques copyrightStatus*, *copyrightJuridiction*, et *copyrightApplicableDates*

5.1 Conclusion sur l'opérationnalisation

Parmi les 23 variables mises en exergue au point 4.6, quatorze se sont vu attribuer des niveaux de qualité permettant la mesure, alors que les neuf autres n'ont pas disposé d'éléments suffisants dans les textes normatifs sélectionnés. Environs 60% de ces variables ont donc un potentiel de mesure au travers d'éléments normatifs identifiés dans les trois normes et standards analysés.

Parmi les variables ne disposant pas de niveaux de qualité, certains éléments appartenant à d'autres normes et standards ont été malgré tout mis en exergue. Le potentiel de mesure et d'automatisation de l'ensemble de ces variables est donc important si l'on s'intéresse à un panorama normatif élargi. Ce dernier point répond à

l'un des objectifs de ce travail, soit la mise en évidence d'éléments permettant une automatisation de la mesure.

6. Conclusion

Les précédents chapitres nous ont permis de définir les concepts clés liés à notre problématique, d'analyser les textes normatifs sélectionnés et d'en extraire les éléments de mesure des dimensions de la qualité des archives, d'organiser ces derniers en un système de mesure, et enfin, de juger de la pertinence et de l'automatisation de ce dernier. A l'issue de ce travail, différentes remarques et suggestions se sont révélées être pertinentes.

Tout d'abord, en ce qui concerne l'organisation des dimensions de la qualité en cadres conceptuels, ou en systèmes de mesure, il s'est avéré qu'il est relativement difficile de modéliser ces derniers en une structure stable, cohérente et universelle en regard des textes analysés. Par exemple, la norme ISO 15489 préconise qu'un document d'activité soit exploitable en le reliant à l'activité ou à l'opération à l'origine de sa création (Organisation internationale de normalisation 2001a, p. 8). Ces liens de contexte de création du document d'activité participent également à la dimension de la fiabilité : "un document fiable est un document dont le contenu peut être considéré comme la représentation complète et exacte des opérations, des activités ou des faits qu'il atteste" (Ibid.). Le contexte de création du document d'activité peut, ici, participer à la mesure de deux dimensions de la qualité des archives. Soit la dimension de l'exploitabilité, via la compréhensibilité, et la dimension de la fiabilité, via des activités ou des faits qui font de lui une "représentation complètes et exactes" de ces dernières.

A un niveau hiérarchiquement inférieur, certains éléments peuvent aussi participer à la mesure de nombreuses dimensions de la qualité des archives. Ainsi, l'*unité sémantique storageMedium* renseigne sur le type de moyen de stockage. Dans ce sens, il est une mesure de la sous-dimension de l'accessibilité technique et de son indicateur de la lisibilité. En effet, afin de planifier la duplication, soit le transfert de l'objet numérique sur un nouveau type de support lors d'obsolescence technologique, cette information est indispensable. Cependant, elle pourrait aussi participer, dans certains cas, à la mesure de la fonctionnalité des accès via ses *composants sémantiques contentLocationType* et *contentLocationValue*. Ces derniers peuvent effectivement permettre de récupérer un objet dans un système de stockage (PREMIS Editorial Committee 2012, p. 76). Dès lors, l'organisation des cadres conceptuels et l'élaboration de variables sont tributaires d'une certaine subjectivité selon que l'on se place dans une approche de la mesure de l'authenticité, de la fiabilité, de l'intégrité, voire de l'exploitabilité. Un des objectifs de ce travail étant d'identifier des éléments de mesure automatisables, une recommandation serait de définir des variables

indépendantes de tous cadres conceptuels. Ainsi, les variables intégrées dans notre opérationnalisation ne se sont pas vues attribuées de références numériques les rattachant à une hiérarchie spécifique. Cette façon de faire devrait permettre leur réutilisation dans d'autres cadres conceptuels spécifiques.

Les normes et standards sélectionnés pour ce travail, soit OAIS, METS et PREMIS ont assurément donnés lieu à la mise en évidence d'éléments participants à la mesure des dimensions de la qualité des archives. Malgré tout, certaines parties du cadre conceptuel de base choisi pour cette analyse n'ont pas pu être mise en exergue dans les textes normatifs. C'est notamment le cas pour la dimension de la représentativité. Cette dernière dimension repose sur des éléments empreints d'une certaine subjectivité. Des éléments de mesure associés à cette dimension ne peuvent donc que difficilement se retrouver dans des textes normatifs. Afin de permettre une opérationnalisation de cette dimension et de ses niveaux inférieurs, une analyse du contexte organisationnel et socio-culturel de l'organisation devrait être mise en œuvre. Le cadre conceptuel ainsi mis en évidence serait donc associé à un contexte spécifique, soit une organisation spécifique, et ne devrait malheureusement pas être exportable sur un large panel d'organisation. Dans le même sens, certaines variables pourraient être optionnelles. Ainsi, la variable de la signature n'est pas forcément requise selon le contexte et la mission d'une institution spécifique.

Une autre remarque concernant les limites de la création de variables et de l'opérationnalisation concerne le champ couvert par les normes et standards sélectionnés. En effet, les chapitres dévolus à l'analyse et à l'opérationnalisation ont mis en évidence la limite des trois textes normatifs choisis dans la mise en exergue d'éléments de mesure et de leur mise en pratique. Ainsi, certaines variables ne sont opérationnalisable qu'à l'aide d'autres contributions normatives. C'est notamment le cas pour l'indicateur de la conformité légale, réglementaire et administrative et ses deux variables. Ces deux variables peuvent être mesurées à l'aide de la norme ISDF. Afin de couvrir l'entier du spectre des dimensions de la qualité des archives électroniques, les normes du Comité international des archives doivent aussi être prises en compte. Ainsi, ISAD(G), ISAAR, ISDF et ISDIAH peuvent grandement contribuer à la mesure de la qualité, au travers notamment de leurs formats d'encodage respectifs. Cette façon de faire peut néanmoins créer des doublons. En effet, certaines variables seraient mesurables à l'aide de plusieurs éléments. C'est le cas entre autre de la variable du versement, qui pourrait être renseigné à l'aide de l'*unité sémantique Ingestion* du dictionnaire PREMIS et de l'élément d'information *Modalités d'entrée* de la norme ISAD(G).

Identifier et éprouver des éléments normalisés permet de garantir l'automatisation de la mesure. En effet, utiliser des éléments reconnaissables sans ambiguïté permet d'exporter le système de mesure sur de nombreuses institutions. La limite de cette automatisation réside sans doute dans la multiplicité des sets de métadonnées. Les éléments identifiés dans ce travail ont été opérationnalisés principalement à l'aide d'*unités sémantique* du dictionnaire de métadonnées PREMIS. Adapter les grilles de mesures à d'autres contextes, soit d'autres types de métadonnées serait un frein éventuel à l'élaboration de méthodes de mesure communes. Malgré tout, l'explicitation des *unités sémantiques* du dictionnaire PREMIS devrait permettre de transposer ces éléments à d'autres schémas normatifs.

Pour finir, ce travail peut constituer une base pour développer des systèmes de mesures adaptés à chaque institution. Il permet en outre d'identifier les éléments indispensables à la préservation des archives électroniques et de juger de leur adéquation à garantir un accès pérenne à des documents électroniques authentiques, fiables et intègres.

Bibliographie

AFNOR, 2012. *ISO 30300 – 30301 – Système de gestion des documents d'activité : définition, modélisation, intégration aux autres normes de système de management*. Livre blanc.3 [en ligne]. Mai 2012. [Consulté le 03.12.2014]. Disponible à l'adresse : http://www.patrimoine.vd.ch/fileadmin/groups/19/PDF/AFNOR-CN11_bilingue.pdf

AN, Xiaomi, 2005. Research in electronic records management. In : MCLEOD, Julie, HARE, Catherine (éd.). *Managing electronic records*. London : Facet Publishing, 2005, pp. 63-80

ARNAUD, Jules, 2012-2014. Gestion des documents d'activité : définition, principes et concepts. In : *Normalisation et gestion des documents d'activité (records management) : enjeux et nouvelles pratiques pour notre profession*. 2012-2014. La Gazette des Archives, n° 228, pp. 59-76. ISSN 0016-5522

BANAT-BERGER, Françoise, DUPLOUY, Laurent, HUC, Claude, 2009. *L'archivage numérique à long terme : les débuts de la maturité ?* Paris : Direction des Archives de France, 2009. Manuels et guides pratiques. ISBN 978-2-11-006942-9

BIBLIOTHEQUE NATIONALE DE FRANCE, 2013. SPAR, le système de préservation numérique de la BnF. *Bibliothèque nationale de France* [en ligne]. 12 novembre 2013. [Consulté le 09.02.2015]. Disponible à l'adresse : http://www.bnf.fr/fr/professionnels/spar_systeme_preservation_numerique.html

BRUNSMANN, Jörg, WILKES, Wolfgang, SCHLAGETER, Gunter, HEMMJE, Matthias, 2012. State-of-the-art of long term preservation in product lifecycle management. *International Journal on Digital Libraries*. 27 janvier 2012

CANTON DE VAUD. GRAND CONSEIL, 2002. *Loi sur l'information (LInfo) du 24 septembre 2002* [en ligne]. 34 septembre 2002. [Consulté le 19.02.2015]. Disponible à l'adresse : http://www.rsv.vd.ch/dire-cocoon/rsv_site/doc.pdf?docId=871790&Pvigueur=&Padoption=&Pcurrent_version=0&PetatDoc=vigueur&Pversion=&docType=loi&page_format=A4_3&isRSV=true&isSJL=true&outformat=pdf&isModifiante=false

CANTON DE VAUD. GRAND CONSEIL, 2011. *Loi sur l'archivage (LArch) du 14 juin 2011* [en ligne]. 14 juin 2011. [Consulté le 19.02.2015]. Disponible à l'adresse : http://www.rsv.vd.ch/dire-cocoon/rsv_site/doc.pdf?docId=871790&Pvigueur=&Padoption=&Pcurrent_version=0&PetatDoc=vigueur&Pversion=&docType=loi&page_format=A4_3&isRSV=true&isSJL=true&outformat=pdf&isModifiante=false

CAYA, Marcel, 2004. La théorie des trois âges en archivistique. En avons-nous toujours besoin ? Éditions en ligne de l'École des chartes (ELEC) [en ligne]. 2 décembre 2004. [Consulté le 09.02.2015]. Disponible à l'adresse : http://elec.enc.sorbonne.fr/conferences/caya

CENTRE DE COORDINATION POUR L'ARCHIVAGE A LONG TERME DES DOCUMENTS ELECTRONIQUES, 2014. Catalogue des formats de données d'archivage. *CECO* [en ligne]. 17 juillet 2014. [Consulté le 21.01.2015]. Disponible à l'adresse : http://www.kost-ceco.ch/wiki/whelp/Cfa/index.php

CENTRE INFORMATIQUE NATIONALE DE L'ENSEIGNEMENT SUPERIEUR, 2014. Le modèle de référence : l'OAIS. *CINES* [en ligne]. 2014. [Consulté le 19.01.2015]. Disponible à l'adresse : https://www.cines.fr/archivage/un-concept-des-problematiques/le-modele-de-reference-loais/

CENTRE NATIONAL DE RESSOURCES TEXTUELLES ET LEXICALES, 2012. *Ortolang. Outils et Ressources pour un Traitement Optimisé de la LANGue* [en ligne].

2012. [Consulté le 08.09.2014]. Disponible à l'adresse : http://www.cnrtl.fr/lexicographie

CENTRE NATIONAL D'ETUDES SPATIALES, 2010. *PAIMAS et PAIS : Producer Interface Methodology Abstract Standard, Producer-Archive Interface Specification* [en ligne]. 2010. [Consulté le 07.02.2015]. Disponible à l'adresse : http://pin.association-aristote.fr/lib/exe/fetch.php/public/presentations/2010/pin20101007-pres04-paimaspais.pdf

CHAUMETTE, Yves, 2006. *La qualité au-delà des mots : perception par la couleur*. Paris : Lavoisier, 2006. Management et informatique. ISBN 2-7462-1494-6

COMITE CONSULTATIF POUR LES SYSTEMES DE DONNEES SPATIALES, 2005. *Modèle de référence pour un Système ouvert d'archivage d'information (OAIS)* [en ligne]. Washington : CCSDS, 2005. Livre Bleu, CCSDS 650.0B-1 (F). Disponible à l'adresse : http://public.ccsds.org/publications/archive/650x0b1s%28F%29.pdf

CONFEDERATION SUISSE. AUTORITES FEDERALES, 2011. *Loi fédérale sur le droit d'auteur et les droits voisins. Loi sur le droit d'auteur, LDA du 9 octobre 1992* [en ligne]. 9 octobre 1992. [Consulté le 19.02.2015]. Disponible à l'adresse : www.admin.ch/opc/fr/classified-compilation/19920251

CONSEIL INTERNATIONAL DES ARCHIVES, 2000. ISAD(G) : *Norme générale et internationale de description archivistique*. 2ème éd. Paris : CIA, 2000. ISBN 0-9696035-6-8

CONSEIL INTERNATIONAL DES ARCHIVES, 2004. ISAAR (CPF) : *Norme Internationale sur les notices d'autorité utilisées pour les Archives relatives aux collectivités, aux personnes ou aux familles*. 2ème éd. Paris : CIA, 2004. ISBN 2-9521932-3-1

CONSEIL INTERNATIONAL DES ARCHIVES, 2007. ISDF : *Norme internationale pour la description des fonctions*. Paris : CIA, 2007

CONSEIL INTERNATIONAL DES ARCHIVES, 2008. ISDIAH : *Norme internationale pour la description des institutions de conservation des archives*. Paris : CIA, 2008

CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS, 2004. *Producer-Archive Interface Methodology Abstract Standard* [en ligne]. Washington : CCSDS, 2004, Magenta Book, CCSDS 651.0-M-1. Disponible à l'adresse : http://public.ccsds.org/publications/archive/651x0m1.pdf

CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEMS, 2012. *Reference Model for an Open Archival Informaion System (OAIS)* [en ligne]. Washington : CCSDS, 2012, Magenta Book, CCSDS 650.0-M-2. Disponible à l'adresse : http://public.ccsds.org/publications/archive/650X0m2.pdf

COTTIN, Michel, DESSOLIN-BAUMANN, Sylvie, 2012-2014. La famille des normes ISO sur le records management. In : *Normalisation et gestion des documents d'activité (records management) : enjeux et nouvelles pratiques pour notre profession*. 2012-2014. La Gazette des Archives, n° 228, pp. 119-133. ISSN 0016-5522

COUNCIL ON LIBRARY AND INFORMATION RESOURCES, 2014. Project Background. *Council on Library and Information Resources* [en ligne]. 2014. [Consulté le 11.02.2015]. Disponible à l'adresse : www.clir.org/pubs/reports/pub87/part1.html

DAPPERT, Angela, ENDERS, Markus, 2010. Digital Preservation Metadata Standards. *National Information Standards Organisation (NISO*) [en ligne]. 2010. Vol 22, Issue 2. [Consulté le 28.01.2015]. Disponible à l'adresse : http://www.loc.gov/standards/premis/FE_Dappert_Enders_MetadataStds_isqv22no2.pdf

DELJANIN, Sandra, 2012. Digital obsolescence. *INFOtheca*. May 2012. N° 1, Volume XIII, pp. 43-53

DELVOSALLE, Christian, 2002. La qualité : des concepts à la pratique. *Pyramides : Revues du Centre d'études et de recherches en administration public* [en ligne]. 2002. N° 5, 2002. pp. 137-154. [Consulté le 03.10.2014]. Disponible à l'adresse : http://pyramides.revues.org/480

DIGITAL CURATION CENTRE, 2004-2015. What is digital curation ? *DCC : because good research needs good data* [en ligne]. 2004-2015. [Consulté le 03.02.2015]. Disponible à l'adresse : www.dcc.ac.uk/digital-curation/what-digital-curation

DIGITAL LIBRARY FEDERATION. *<METS> Metadata Encoding and Transmission Standard : Primer and Reference Manual* [en ligne]. [s.l.] : Digital Library Federation, 2010. Version 1.6 Revised. [Consulté le 09.02.2015]. Disponible à l'adresse : http://www.loc.gov/standards/mets/METSPrimerRevised.pdf

DIGITAL PRESERVATION COALITION, 2012. Introduction - Definitions and Concepts. *Digital Preservation Coalition* [en ligne]. 2012. [Consulté le 16.01.2015]. Disponible à l'adresse : http://www.dpconline.org/advice/preservationhandbook/introduction/definitions-and-concepts

DOYLE, Julie, VIKTOR, Herna, PAQUET, Eric. Long-term digital preservation : preserving authenticity and usability of 3D-data. *Journal on Digital Libraries*. 2009-03. pp. 33-47

DUBOIS, Alain, WILDI, Tobias, 2012. Archivage électronique - Matterhorn METS Profile enregistré auprès de la Bibliothèque du Congrès. *ARBIDO* [en ligne]. 03.12.12. [Consulté le 09.02.2015]. Disponible à l'adresse : http://www.arbido.ch/fr/artikel_detail.php?m=3&id=1556&n=121

DUBOIS, Alain, WILDI, Tobias, 2013. Vers une intégration complète des standards de description du Conseil international des Archives ? Réflexions autour du processus décisionnel du Conseil d'Etat valaisan. *Slideshare* [en ligne]. 20 juin 2013. [Consulté le 25.01.2015]. Disponible à l'adresse : http://fr.slideshare.net/TobiasWildi/presentation-geneve-20130617#

DUNANT GONZENBACH, Anouk, DUCRY, Emmanuel, 2013. *L'archivage des documents électroniques à Genève : aspects organisationnels et techniques. Le projet Gal@tae* [en ligne]. Décembre 2013. [Consulté le 25.02.2015]. Disponible à l'adresse : http://hieretdemain.gonzen.com/2013_galatae_blog.pdf

FAIVRE D'ARCIER, Louis, 2004. La collecte des archives électroniques aux Archives de Paris : quelques exemples d'interventions des archivistes. In : FILLIEUX, Véronique, VANDEVOORDE, Evelyne. *Les archives électroniques : quels défis pour l'avenir ? : actes de la Troisième Journée des Archives, Louvain, 8 et 9 mai 2003*. Louvain-la-Neuve : Bruylant-Academia, 2004, pp. 51-69. Publication des Archives de l'Université Catholique de Louvain, 8. ISBN 2-87209-757-0

FLYNN, Sarah J. A., 2001. The Records Continuum Model in Context and its Implication for Archival Practice. *Journal of the Society of Archivists*. 2001. Vol 22, No. 1, pp. 79-93

FROMAN, Bernard, GOURDON, Christophe, 2003. *Dictionnaire de la qualité*. Saint-Denis-La-Plaine : AFNOR, 2003. ISBN 2-12-467821-3

GUENTHER, Rebecca, XIE, Zhiwu, 2007. *Implementing PREMIS in Container Formats* [en ligne]. 2007. [Consulté le 28.01.2015]. Disponible à l'adresse : www.loc.gov/standards/premis/ist-containers.ppt

HARVEY, Ross, 2010. *Digital curation : a how-to-do-it manual*. London : Facet Publishing, 2010. ISBN 978-1-85604-733-3

HOFMAN, Hans, 2005. The Use of Standards and Models. In : MCLEOD, Julie, HARE, Catherine (éd.). *Managing electronic records*. London : Facet Publishing, 2005, pp. 18-33.

INTERPARES 2 PROJECT, 2012. Terminology Database. *International Research on Permanent Authentic Records in Electronic Systems* [en ligne]. 2012. [Consulté le 08.12.2014]. Disponible à l'adresse : http://www.interpares.org/ip2/ip2_terminology_db.cfm

JACCARD, Michel, 2010. *Objectif qualité : introduction aux systèmes de management de performance et de durabilité*. Lausanne : Presses polytechniques et universitaires romandes, 2010. ISBN 978-2-88074-867-8

KALBERMATTEN, Brigitte, 2012. Les métadonnées dans les paquets d'information numériques SIP et AIP: de leur création à leur gestion. In : COUTAZ, Gilbert, KNOCH-MUND, Gaby, REIMER, Ulrich (éd.). *Informationswissenschaft : Theorie, Methode und Praxis = Sciences de l'information : théorie, méthode et pratique*. Travaux du/Arbeiten aus dem Master of Advanced Studies in Archival, Library and Information Science, 2010-2012. Baden : Hier und Jetzt, 2014, pp. 33-46

KENT, Allen (Ed.), 2002. *Encyclopedia of library and information science*. New-York : Marcel Dekker, 2002. Vol. 71. ISBN 0-8247-2072-5

LIBRARY OF CONGRESS, 2011. *METS : vue d'ensemble & tutoriel* [en ligne]. 1er juillet 2011. [Consulté le 16.10.2014. Disponible à l'adresse : www.loc.gov/standards/mets/METSOverview.v2_fr.html

LIBRARY OF CONGRESS NETWORK DEVELOPMENT AND MARC STANDARDS OFFICE, 2009. *Comprendre PREMIS* [en ligne]. 1er février 2009. [Consulté le 03.11.2014]. Disponible à l'adresse : http://www.loc.gov/standards/premis/Understanding-PREMIS_french.pdf

MAKHLOUF SHABOU, Basma, 2010. *Etude sur la définition et la mesure des qualités des archives définitives issues d'une évaluation*. Montréal : Université de Montréal. Thèse présenté à la Faculté des études supérieures en vue de l'obtention du grade de Philosophiae Doctor (Ph.D) en science de l'information, Université de Montréal, 2010

MAKHLOUF SHABOU, Basma, 2013. Le projet QADEPs : un outil au service de la pérennisation des données publiques. In : HIRAUX, Françoise, MIRGUET, Françoise. *De la préservation à la conservation : stratégies pratiques d'archivage : actes des Journées 2013 des Archives de l'UCL, Louvain, 18 et 19 avril 2013*. Louvain-la-Neuve : Academia-L'Harmattan, 2014, pp. 87-98. Publication des Archives de l'Université Catholique de Louvain. ISBN 978-2-8061-0166-2

NATIONAL INFORMATION STANDARDS ORGANISATION. Planets Survey Gauges Organizational Readiness for Digital Preservation. *Information Standards Quarterly*. Spring 2010, Vol. 22, Issue 2

NEF, Andreas, WILDI, Tobias, 2013. *Digital archivieren mit CMI STAR und Documill* [en ligne]. 24.4.2013. [Consulté le 09.02.2015]. Disponible à l'adresse : www.cmistar.ch/.../Tobias_Wildi_-_Live_OAIS.pdf

OFFICE QUEBECOIS DE LA LANGUE FRANCAISE, 2012. Le grand dictionnaire terminologique. *Office québécois de la langue française* [en ligne]. 2012. [Consulté le 08.09.2014]. Disponible à l'adresse : http://www.granddictionnaire.com/

ORGANISATION INTERNATIONALE DE NORMALISATION, 2001. *Information et documentation – "Records management" – partie 1 : principes directeurs*. Genève : ISO, 2001. ISO 15489 - 1

ORGANISATION INTERNATIONALE DE NORMALISATION, 2001. *Information et documentation – "Records management" – partie 2 : guide pratique*. Genève : ISO, 2001. ISO 15489 – 2

ORGANISATION INTERNATIONALE DE NORMALISATION, 2005. *Système de management de la qualité – principes essentiels et vocabulaire*. Genève : ISO, 2008. ISO 9000 : 2005(F)

ORGANISATION INTERNATIONALE DE NORMALISATION, 2011. *Information et documentation – Système de gestion des documents d'activité – Principes essentiels et vocabulaire*. Genève : ISO, 2011. ISO 30300 : 2011(F)

ORGANISATION INTERNATIONALE DE NORMALISATION, 2011. *Information et documentation – Système de gestion des documents d'activité – Exigences*. Genève : ISO, 2011. ISO 30301 : 2011(F)

PREMIS EDITORIAL COMMITTEE. *PREMIS Data Dictionary for Preservation Metadata* [en ligne]. [s.l.] : PREMIS Editorial Committee, 2012. Version 2.2. [Consulté le 09.02.2015]. Disponible à l'adresse : www.loc.gov/standards/premis/v2/premis-2-2.pdf

PRESERVATION AND LONG-TERM ACCESS THROUGH NETWORKED SERVICES. Welcome to Planets. *Planets* [en ligne]. 2007. [Consulté le 09.02.2015]. Disponible à l'adresse : http://www.planets-project.eu/

ROUSSEL, Stéphanie, 2012-2014. Le champ normatif de l'archivage électronique. In: *Normalisation et gestion des documents d'activité (records management) : enjeux et nouvelles pratiques pour notre profession*. 2012-2014. La Gazette des Archives, n° 228, pp. 59-76. ISSN 0016-5522

SCHELLINGS, Franck, 2004. Développer un système de gestion globale des archives. Le projet DISSCO et les aspects humains. In : FILLIEUX, Véronique, VANDEVOORDE, Evelyne. *Les archives électroniques : quels défis pour l'avenir ? : actes de la Troisième Journée des Archives, Louvain, 8 et 9 mai 2003*. Louvain-la-Neuve : Bruylant-Academia, 2004, pp. 37-49. Publication des Archives de l'Université Catholique de Louvain, 8. ISBN 2-87209-757-0

SOCIETY OF AMERICAN ARCHIVISTS AND THE NETWORK DEVELOPMENT AND MARC STANDARDS OFFICE OF THE LIBRARY OF CONGRESS, 2002. *Encoded Archival Description Tag Library* [en ligne]. [Consulté le 24.01.2015]. Disponible à l'adresse : http://www2.archivists.org/sites/all/files/EAD2002TL_5-03-V2.pdf

SOCIETY OF AMERICAN ARCHIVISTS, STAATSBIBLIOTHEK ZU BERLIN. Welcome to EAC-CPF Homepage. *Encoded Archival Context Corporate Bodies, Persons, and Families* [en ligne]. [Consulté le 24.01.2015]. Disponible à l'adresse : http://eac.staatsbibliothek-berlin.de/

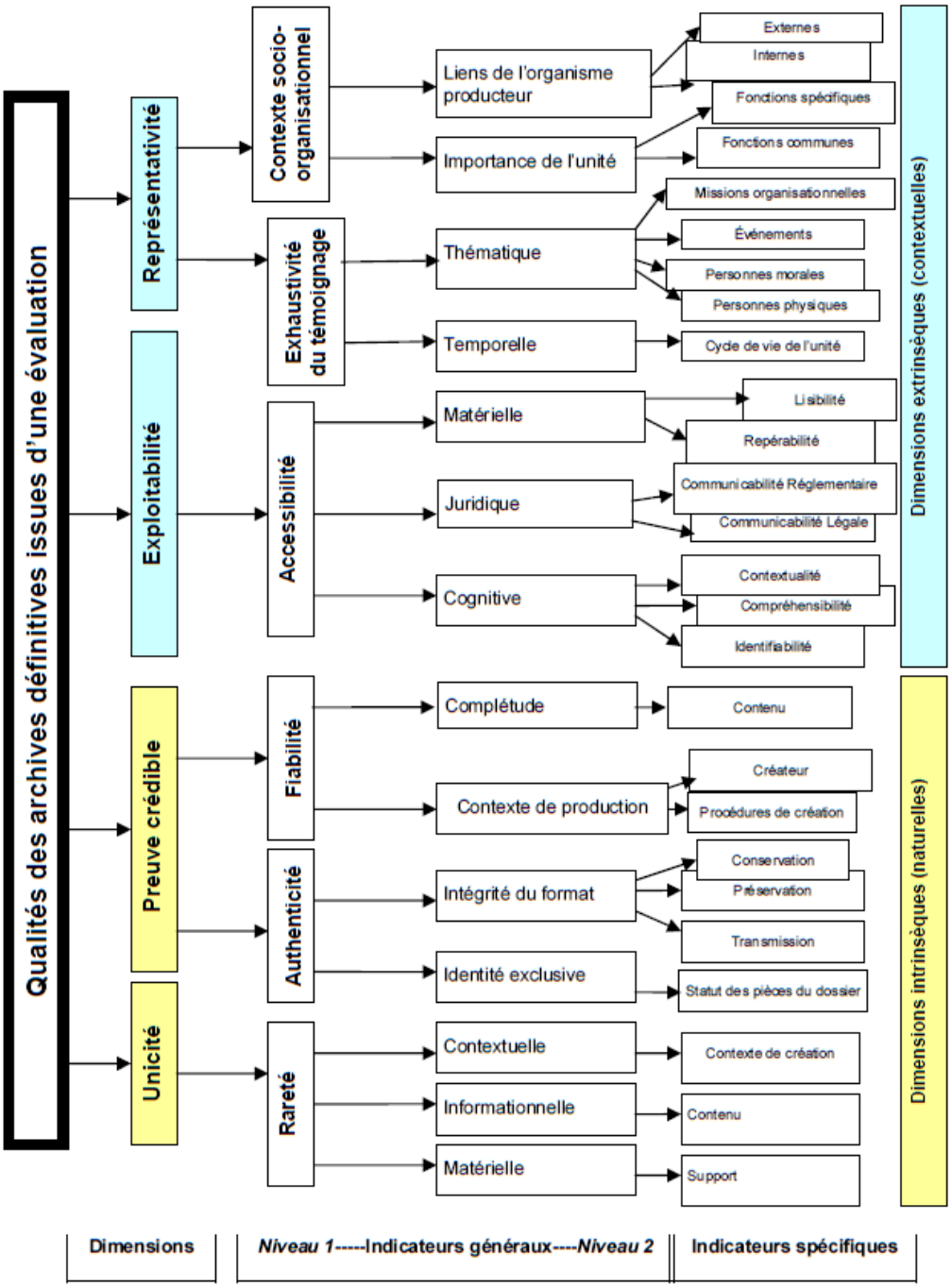
SVARD, Proscovia, 2013. Entreprise Content Management and the Records Continuum Model as strategies for long-term preservation of digital information. *Records Management Journal*. 2013. Vol. 23, No 3, pp. 159-176

TRACE, Ciaran, SANETT, Shelby, 2000. InterPARES : Securing the Future of Our Electronic Records. *Bulletin of the American Society for Information Science*. October / November 2000. pp 24-26

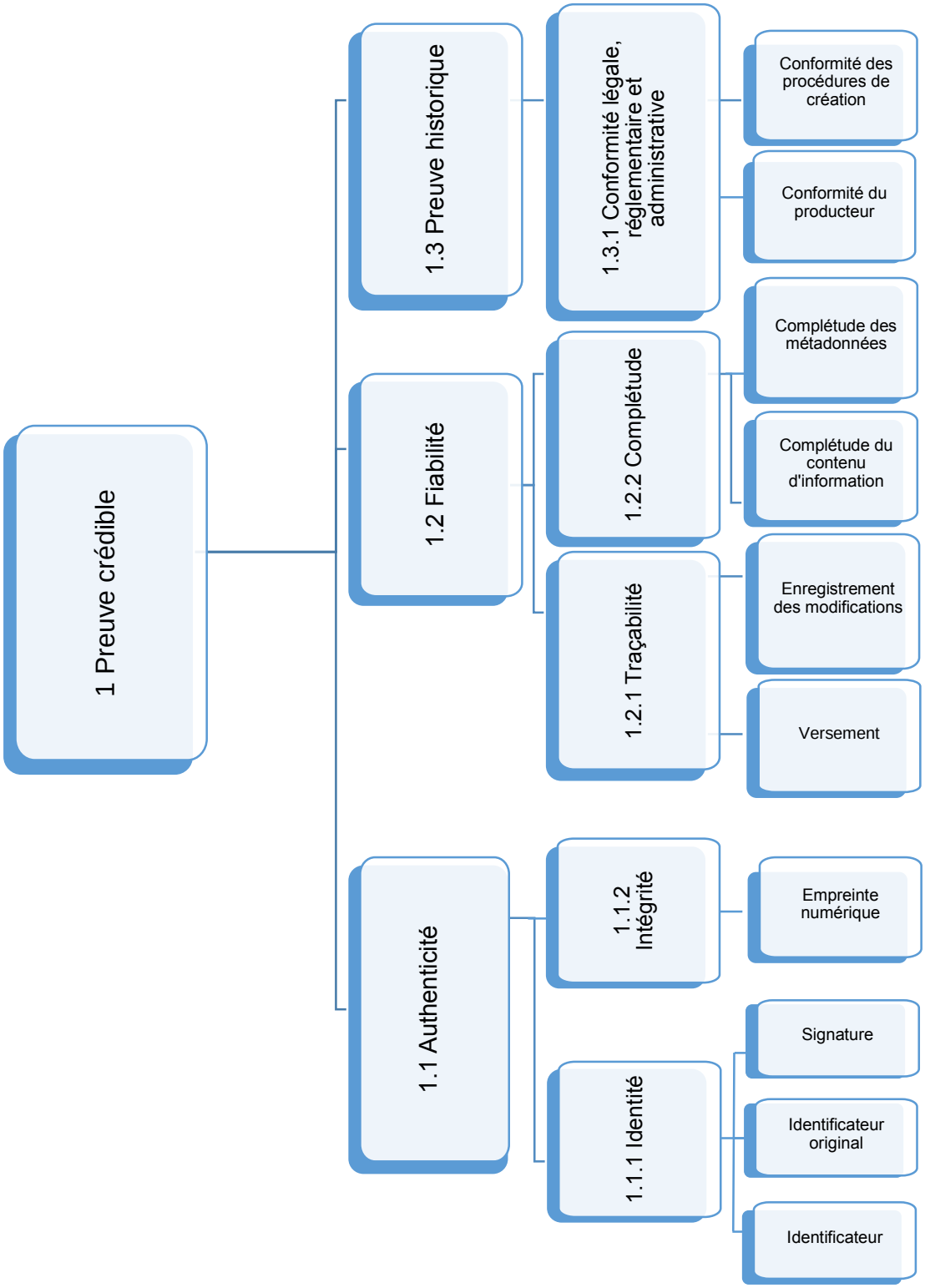
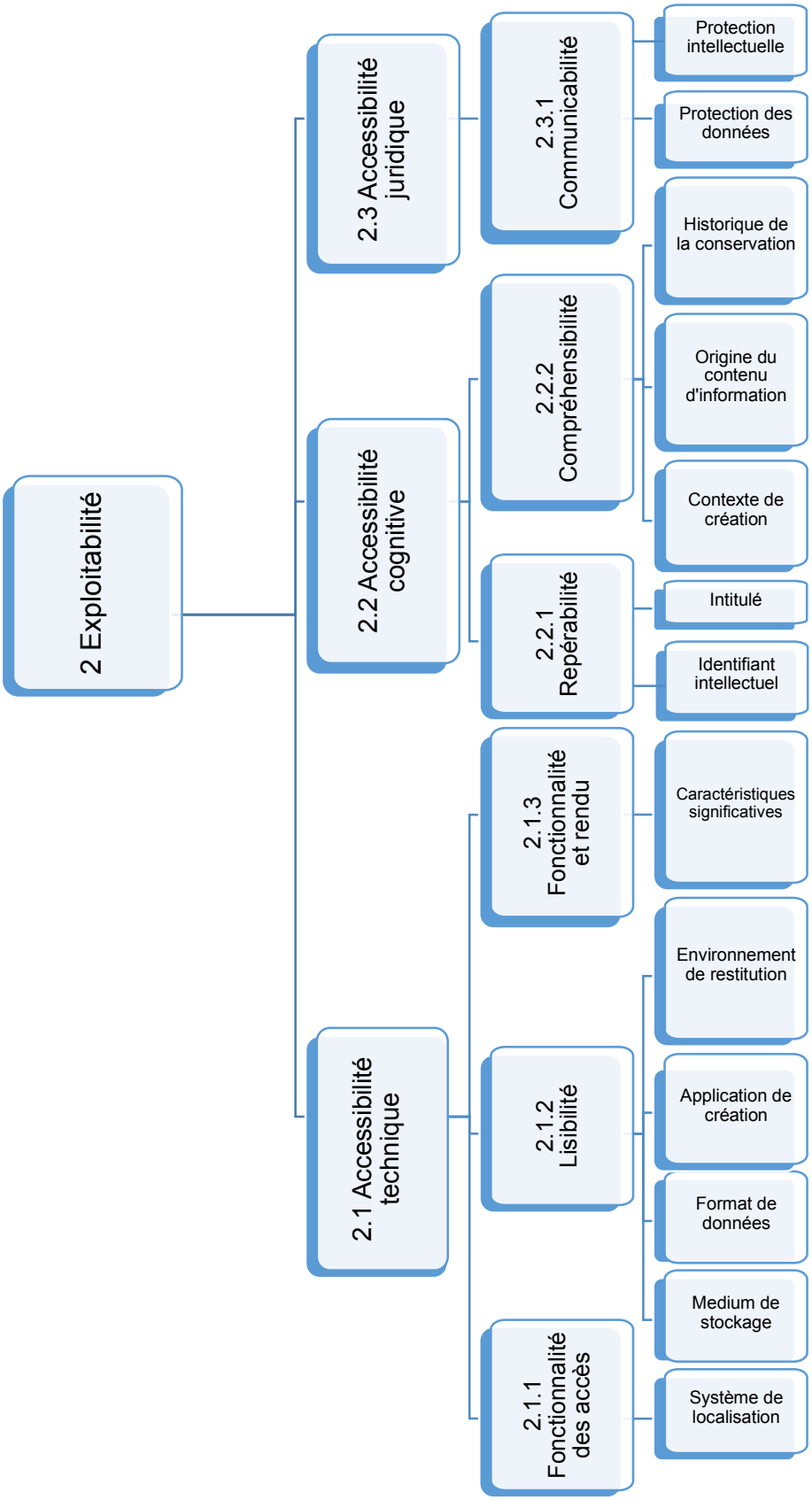
WILDI, Tobias, 2012. Normen und Standards als Synergiepotenziale in der digitalen Archivierung. *ARBIDO*. 14 juin 2012. N° 2, pp. 13-15

WILSON, Andrew, 2008. Significant Properties of Digital Objects [en ligne]. [Consulté le 15.02.2015]. Disponible à l'adresse : http://www.dpconline.org/docs/events/080407sigpropsWilson.pdf

Annexe 1 : cadre conceptuel *étude sur la définition et la mesure des qualités des archives définitives issues d'une évaluation*



Annexe 2 : cadre conceptuel de la preuve crédible



Annexe 3 : cadre conceptuel de l'exploitabilité