

Diplomatique et qualités des documents et données
électroniques publics : quelle leçon retenir?
Aperçu du projet QADEPs 2013

Par Basma Makhoulf Shabou, PhD
Professeure à la Haute école de gestion de Genève
Département Information documentaire

Digital Diplomats 2013: Paris, 14-16 novembre 2013

Introduction

Projet QADEPs

équipe et partenaire

objectifs

approche méthodologie

fondements

aperçu des résultats

Un *focus* sur l'authenticité

exemple de mesures

présentation du guide d'application

possibilité automatisation

Conclusion

Diplomatique et qualités des archives: quels liens?

h e g

Haute école de gestion
Genève

La Diplomatique numérique (DN) : l'examen des caractéristiques des documents numériques afin de **valider** / attester leur fonction de **preuve**

(Duranti, 2009)

DN traite essentiellement des **caractéristiques intrinsèques** des documents et des archives numériques

L'étude de certaines qualités des archives numériques s'appuie fortement sur les méthodes de la DN.

Qualités des archives: Précisions terminologiques

h e g

Haute école de gestion
Genève

Nous comprenons les qualités des archives comme étant les caractéristiques souhaitées dans des documents/ données qui dans leur ensemble permettent d'assurer adéquatement la **constitution d'un patrimoine** documentaire significatif et **d'augmenter leur potentiel d'utilisation**

(Makhlouf Shabou, 2011, pp.8-9)

Types de qualités

Qualités intrinsèques, caractères naturels des documents

Exemple : Unicité et caractère de preuve crédible, etc.

(Trustworthiness, InterPares, 2013)

Qualités extrinsèques : caractères contextuels des documents

Exemple : Exploitabilité, Représentativité, etc.

Définition et mesure des qualités des archives et documents électroniques publics (QADEPs)

Equipe ...

Basma Makhoulf Shabou, Prof. cheffe de projet

Raphaël Rey, Assistant HES

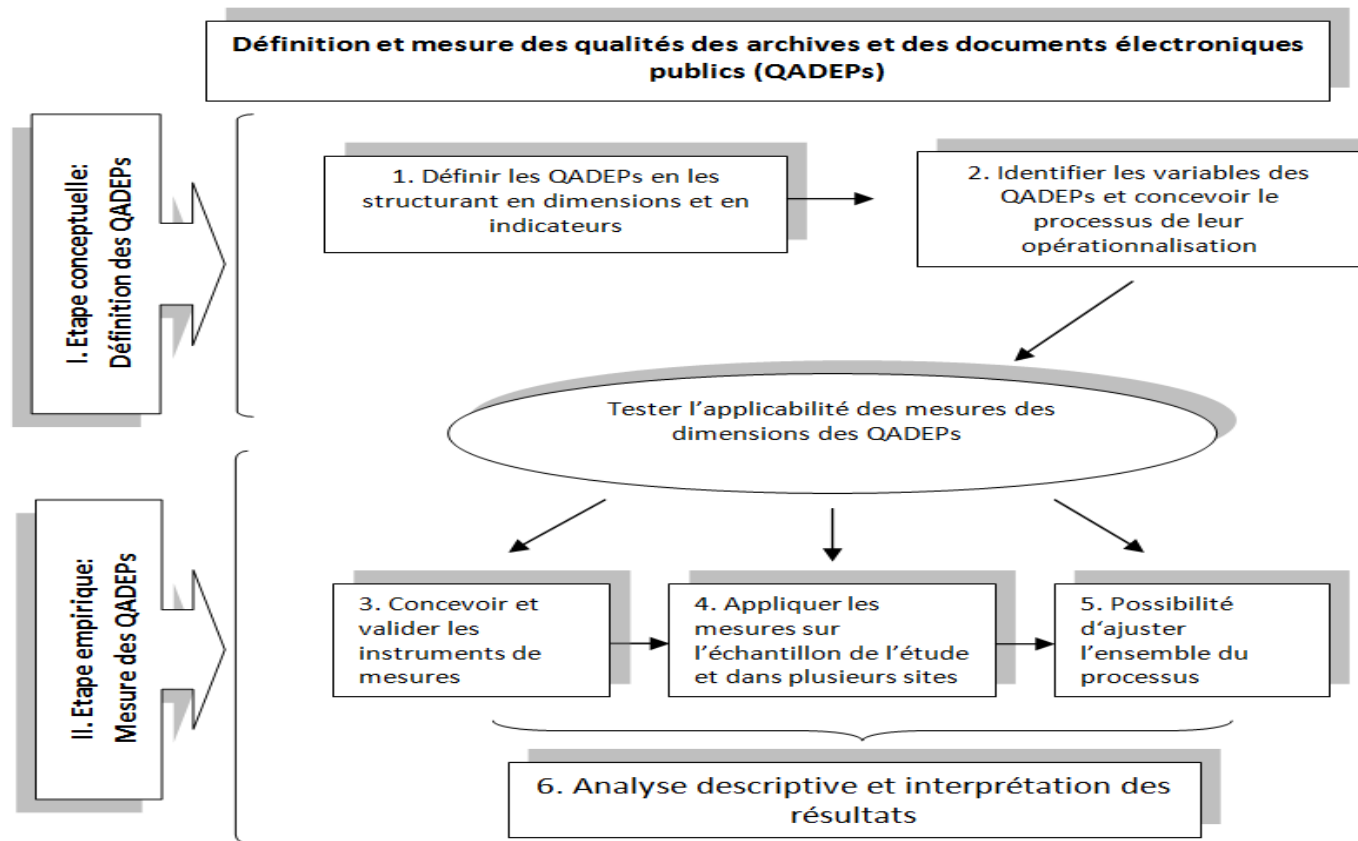
Avec la participation de ...

- Alexandre Boder, Chargé de cours
- Dr René Schneider, Prof HES
- Mellifluo Laure, Assistante HES

Partenaires externes ...

- **Privé** : Docuteam Sàrl, Dr. Tobias Wildi
- **Publics** : Archives fédérales suisses, M. Andreas Kellerhals
 - Archives de l'État du Valais, M. Alain Dubois
 - Archives d'État de Genève, M. Pierre Flückiger
 - Institut d'Histoire de l'Université de Berne, Dr. Christophe Graf

1. Définir les principales **dimensions de qualité** des archives et documents électroniques publics (QADEPs) en identifiant les **indicateurs** qui leur sont appropriés
2. Identifier les **variables** et définir la **démarche opérationnelle** permettant d'évaluer le niveau de qualité des documents électroniques
3. **Tester les mesures** des QADEPs publics afin de vérifier leur **applicabilité** et leur **automatisation**



Fondements

- Thèses et travaux académiques
- Normes et standards (Digital Repositories, RM, métadonnées)

Autres sources

- Recommandations et *best practices* (archives nationales, régionales, cantonales)
- Projets internationaux (InterPares, etc.)
- Revues et conférences professionnelles

Echantillon et périmètre

- archives, documents et données électroniques
- 2 objets y compris des SIP choisis dans deux sites : AGE et le AEV

QADEPs: synthèse des résultats conceptuels

h e g

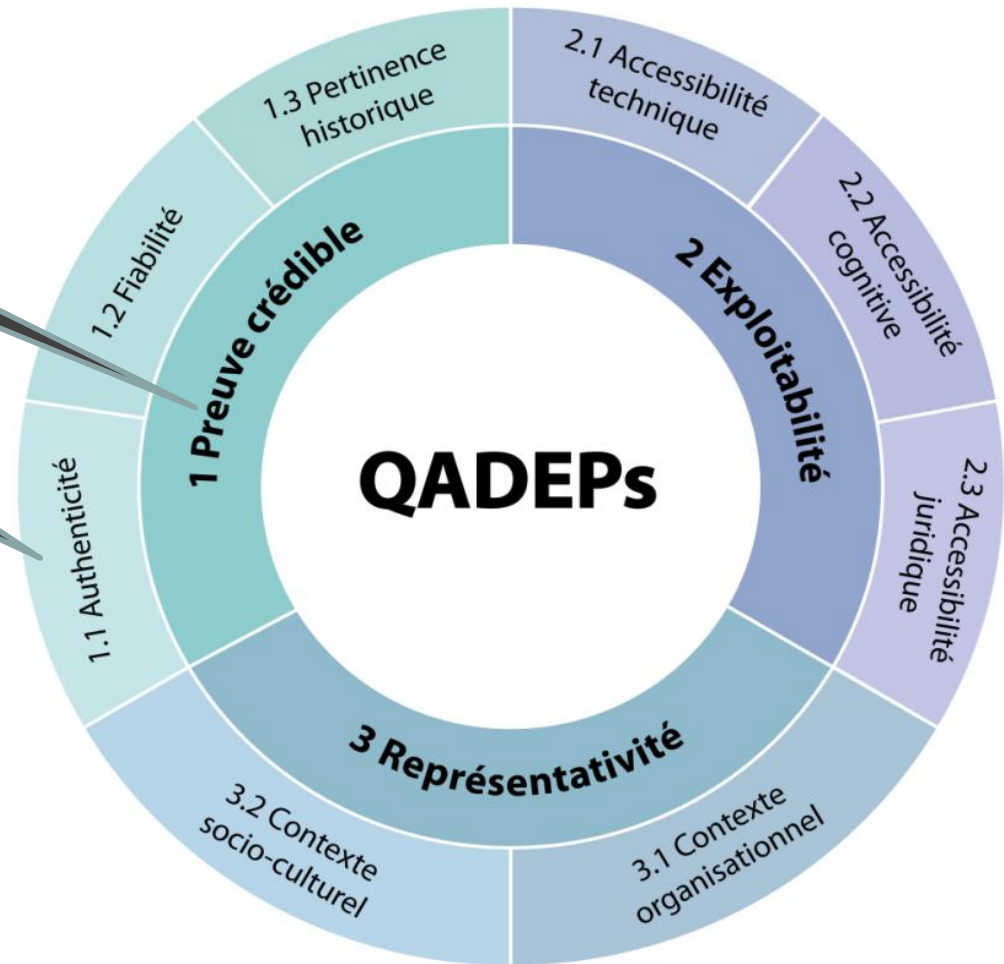
Haute école de gestion
Genève

3
Dimensions

8
Sous -dimensions



46 variables

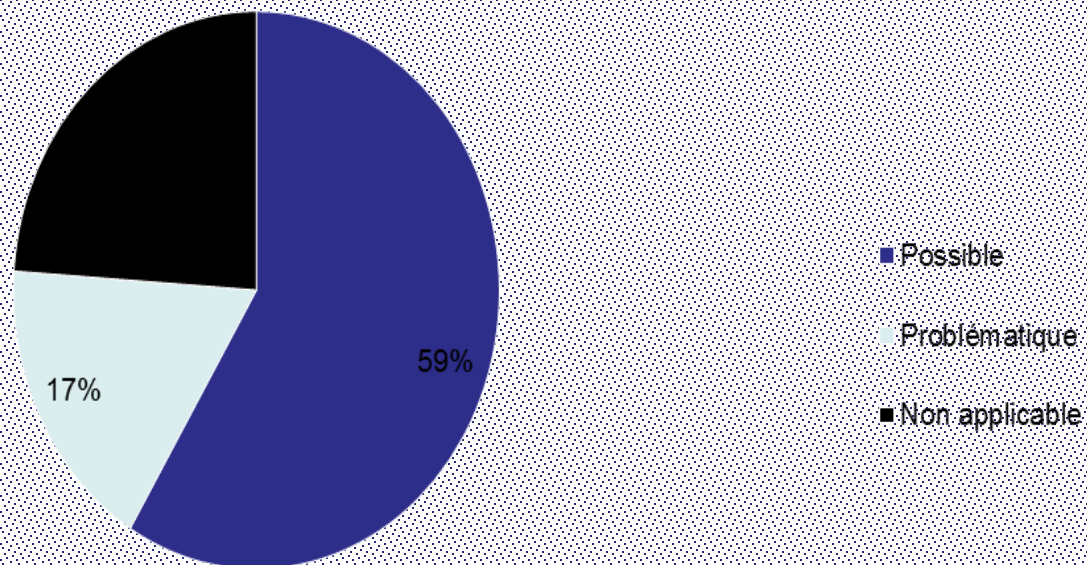


Quatre sous-objectifs ont guidé l'identification de ces variables :

- favoriser le caractère automatisable → un avantage indéniable pour le traitement de données abondantes ;
- limiter le caractère subjectif des évaluations autant que possible ;
- documenter l'application des variables afin de permettre la reproductibilité des mesures ;
- confirmer l'applicabilité et la pertinence des variables via des tests dans plusieurs institutions et sur plusieurs types d'objets informationnels.

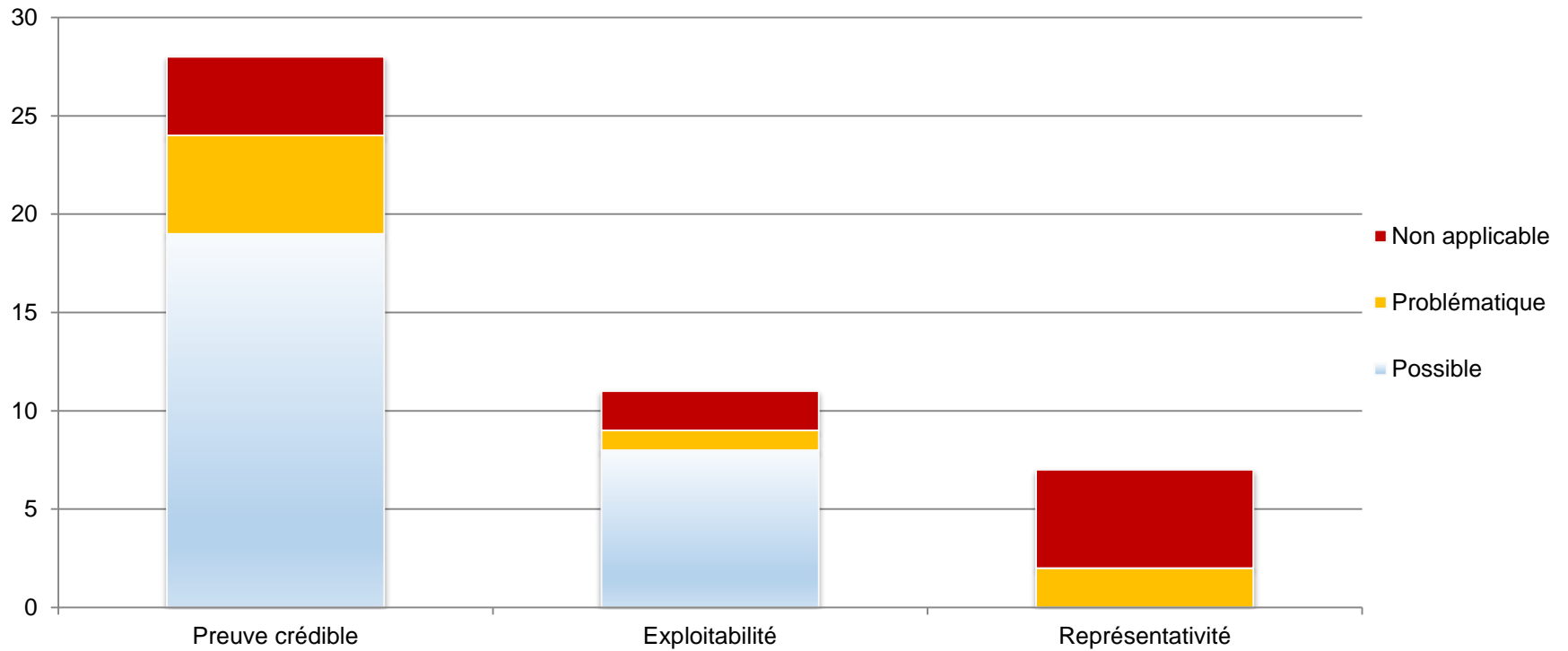
Aperçu sommaire de l'automatisation des variables testées

Automatisation des mesures

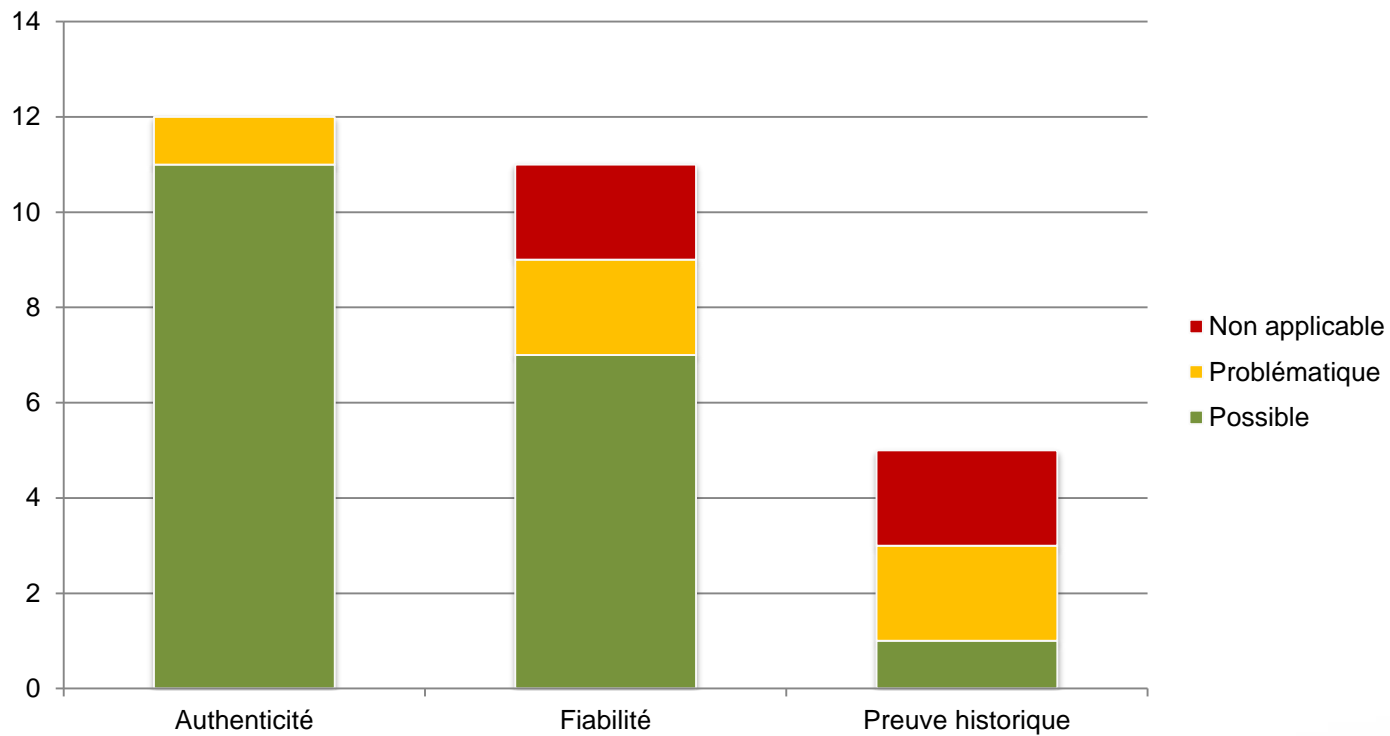


Automatisation des mesures

Répartition par dimension



Automatisation des mesures Preuve crédible



Un *Focus* sur l'authenticité... exemple de mesures

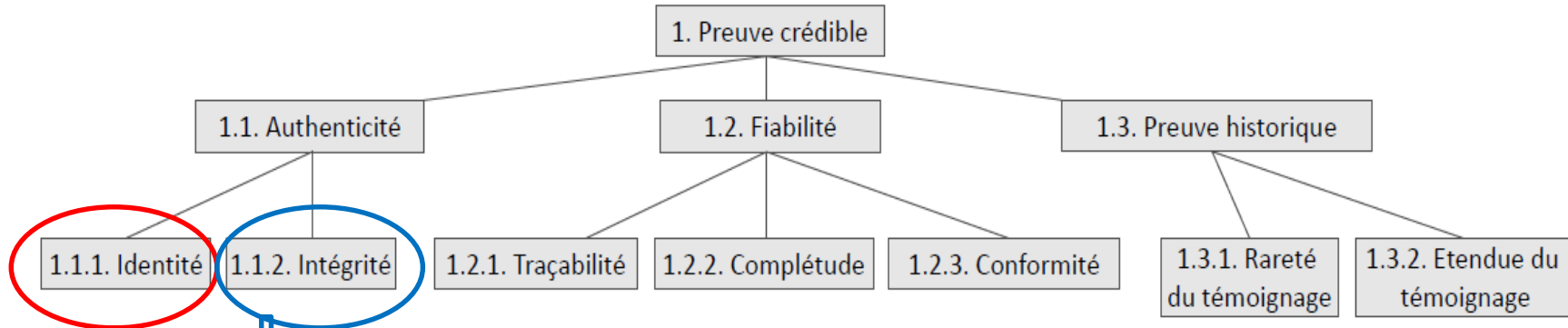
h e g

Haute école de gestion
Genève

Dimension

Sous-
Dimensions

Indicateurs



Variables		Niveaux de qualité
v. 1	Existence d'un identifiant	1. Pas d'identifiant existant 2. Identifiant existant mais non standardisé 3. Identifiant existant standardisé à l'intérieur de l'institution 4. Identifiant existant standardisé à l'intérieur et à l'extérieur de l'institution
...		
v. 12	Fixité de la chaîne de bits	1. Fixité de la chaîne de bits garantie par aucun contrôle 2. Fixité de la chaîne de bits garantie par un contrôle via checksum 3. Fixité de la chaîne de bits garantie par un contrôle périodique via checksum

Un Focus sur l'authenticité... Guide d'application des mesures

h e g

Définitions

B. Descriptif détaillé des variables

1. Dimension : preuve crédible

Définition : capacité d'un document à gagner la confiance de son utilisateur comme étant la source privilégiée à l'appui des faits. Cette qualité repose sur l'authenticité et la fiabilité. (MAKHOLOUF SHABOU, 2010, p. 115 ; INTERPARES 2, 2013)

1.1. Sous-dimension : authenticité

Définition : « un document authentique est un document dont on peut prouver

- qu'il est bien ce qu'il prétend être,
- qu'il a été effectivement produit ou reçu par la personne qui prétend l'avoir produit ou reçu, et
- qu'il a été produit ou reçu au moment où il prétend l'avoir été. »

(ISO 15489-1:2001, p. 7)

1.1.1. Indicateur : identité

Définition : « identity :The whole of the characteristics of a document or a record that uniquely identify it and distinguish it from any other document or record. With integrity, a component of authenticity. » (INTERPARES 2, 2013)

v1. Existence d'un identifiant

- pas d'identifiant existant
- identifiant existant mais non standardisé
- identifiant existant et standardisé à l'intérieur de l'institution
- identifiant existant et standardisé à l'intérieur et à l'extérieur de l'institution

Note d'application : cette variable examine l'existence, l'unicité, la standardisation et la pérennisation des identifiants.

Niveau 1 : s'applique lorsque l'objet ne possède pas d'identifiant désigné en tant que tel.

Niveau 2 : s'applique lorsqu'un identifiant existe, mais qu'il n'y a pas de règles explicites fixant son établissement. L'unicité n'est donc pas absolument garantie. Ce niveau accueille, dans la plupart des cas, les URI étant donné que leur pérennité n'est pas toujours garantie.

Niveau 3 : s'applique lorsque les identifiants sont déterminés par des règles propres à l'institution. L'unicité est garantie, mais uniquement au sein de l'institution. On pourrait par contre trouver le même identifiant attribué à un autre objet dans une institution différente. A noter que les identifiants recourant aux codes ISIL attribués par un centre national appartiennent en général à ce niveau parce que le code ISIL n'enregistre que l'institution. Le reste de l'identifiant est géré et enregistré à l'interne.

Niveau 4 : s'applique lorsque les identifiants sont déterminés selon des normes ou des standards et sont enregistrés dans des registres externes à l'institution. Ces identifiants doivent être attribués de manière définitive et leur unicité garantie. Leur pérennité doit être assurée quelles que soient les modifications ou déplacements que pourrait subir l'objet. Exemple : *Digital object identifier (DOI)*¹, *Persistent URLs (PURL)*², etc.

Remarques : un même objet peut recevoir plusieurs identifiants : par exemple, l'un pour le fonctionnement interne de l'institution et un autre externe de type DOI par exemple. L'évaluation de la variable tiendra compte uniquement de l'identifiant qui présente la valeur la plus élevée.

Sources :

ORGANISATION INTERNATIONALE DE NORMALISATION. *ISO 15511 : Information et documentation - Identifiant international normalisé pour les bibliothèques et les organismes apparentés (ISIL)*. Genève. 2011.

¹Les DOI se composent d'un préfixe et d'un suffixe de la manière suivante : 10.xxxxxx/xxxxxxx. Le « 10. » est propre à tous les DOI. Le préfixe sert à désigner l'agence qui enregistre les DOI, quant au suffixe, il permet d'identifier l'objet lui-même au sein du registre. Le suffixe peut prendre des formes très diverses comme un ISSN ou n'importe quelle chaîne de caractères.

²cf. PURL'S CONTRIBUTORS, 2011 : « PURLs are Web addresses or Uniform Resource Locators (URLs) that act as permanent identifiers in the face of a dynamic and changing Web infrastructure. Instead of resolving directly to Web resources (documents, data, services, people, etc.) PURLs provide a level of indirection that allows the underlying Web addresses of resources to change over time without negatively affecting systems that depend on them. This capability provides continuity of references to network resources that may migrate from machine to machine for business, social or technical reasons. »

Notes d'application

Remarques

Sources

Niveaux de qualité

La mesure des QADEPs est...

- possible
- vérifiable
- et en grande partie automatisable

La méthode proposée par cette étude

- est reproductible
- permet des résultats qui ont le pouvoir discriminatoire

Ce que nous retenons aussi...

- Le caractère multidimensionnel des QADEPs
- Les résultats issus de la mesure des QADEPs sont relatifs

Conceptuelles

- Offrir une définition au concept de qualité adapté aux archives électroniques

Professionnelles

- Offrir un instrument de mesure et de contrôle des QADs;
- Consolider la justification des acquisitions;

Méthodologique

- Proposer une méthode testée, reproductible et vérifiable de mesure des qualités applicables aux archives avec les dispositifs de sa réalisation

Académiques

- Intégration du projet InterPares Trust et RiC (2013)
- Travail de Bachelor ID (2014)
- Un gp d'étudiants de la Filière Informatique de gestion travaille sur le développement d'un software QADEPs...

Merci de votre attention !

basma.makhlouf-shabou@hesge.ch