



HAL
open science

Evaluer la qualité des archives ouvertes : le certificat DINI

Joachim Schöpfel, Uwe Müller

► **To cite this version:**

Joachim Schöpfel, Uwe Müller. Evaluer la qualité des archives ouvertes : le certificat DINI. Partnership, 2014, 9 (1), 10.21083/partnership.v9i1.2733 . hal-01226199

HAL Id: hal-01226199

<https://hal.univ-lille.fr/hal-01226199>

Submitted on 9 Nov 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Evaluer la qualité des archives ouvertes : le certificat DINI

Joachim Schöpfel
Université de Lille 3
joachim.schopfel@univ-lille3.fr

Uwe Müller
Deutsche Nationalbibliothek
u.mueller@dnb.de

Résumé

L'article présente le certificat DINI, un référentiel pour la certification et l'audit des archives ouvertes, des archives institutionnelles et des plateformes de revues en libre accès, développé par l'Initiative Allemande pour l'Information en Réseau DINI. L'article décrit le contexte, l'objectif et l'historique de ce certificat avant d'exposer sa structure et son contenu. Parmi les huit sections du certificat figurent la visibilité du site, la sécurité de l'information et l'archivage pérenne. La discussion porte sur l'objet du référentiel, sur son influence et sur la nécessité d'une adaptation au contexte francophone. La traduction française du certificat a été publiée en 2012.

Abstract

The article presents the DINI certificate, standard recommendations for the certification of document and publishing services, i.e. open archives, institutional repositories and platforms for open access journals, developed by the German Initiative for Networked Information DINI. The article describes the context, purpose and history of this certificate before exposing its structure and content. The eight areas include criteria for the visibility of the site, information security and long-term preservation. The article discusses the object of certification, its impact and the need to adapt some recommendations to the French-speaking context. The French translation of the certificate has been published in 2012.

Mots-clés

archives ouvertes; archives institutionnelles; libre accès; certification; normalisation; qualité; évaluation

Keywords

open archives; institutional repositories; open access; certification; standardization; quality; evaluation

Introduction

Depuis vingt ans, le mouvement vers le libre accès (*open access*) modifie en profondeur le système de l'information scientifique¹. Au départ, l'idée du libre accès reflétait surtout le concept de la communication scientifique directe, sur le modèle de la plateforme arXiv de la communauté des *High Energy Physics* (physique des particules), considérée comme la première archive ouverte². L'idée des chercheurs était simple : s'approprier des technologies du Web pour diffuser sans délai au sein de la communauté scientifique les manuscrits (*preprints*) envoyés aux revues scientifiques, avant toute sélection, révision (*peer review*) et publication formelle. Autrement dit, chaque auteur dépose lui-même son manuscrit sur une plateforme pour le stockage (auto-archivage) et la diffusion sans restriction sur le Web. Cette idée incarnait aussi la réaction face à l'augmentation rapide des coûts d'abonnements aux revues scientifiques et aux barrières économiques de la diffusion des résultats de la recherche. Le site d'auto-archivage le plus connu dans les pays francophones est la plateforme HAL du Centre pour la Communication Scientifique Directe (CCSD) du CNRS en France, proposant en 2013 plus de 230 000 documents avec texte intégral – surtout des articles mais aussi des thèses, rapports, communications etc.³

En parallèle, sont apparues les premières revues en libre accès, telles que *Applied Semiotics/Sémiotique appliquée*, *Belphégor* ou *VertigO*, puis dans un second temps des serveurs pour l'édition des revues en libre accès, tels que *Érudit* au Canada⁴, *Revue.org* en France⁵ ou *SciELO* au Brésil⁶ qui totalisent ensemble plus d'un millier de titres avec plusieurs centaines de milliers d'articles en libre accès. Là encore, l'idée initiale était d'utiliser les nouvelles technologies pour développer la fonction éditoriale au sein des communautés, équipes et établissements scientifiques et pour diffuser les articles sélectionnés sur Internet (Web) aux lecteurs intéressés, sans licences ni coûts d'abonnements. Aujourd'hui, l'idée initiale s'est transformée en un nouveau marché où cohabitent des initiatives publiques, parfois avec de nouveaux modèles économiques⁷, et

¹ Voir Suber 2012 pour une vue globale du mouvement et le site de veille « Libre accès à l'information scientifique et technique » de l'INIST (CNRS) pour les actualités, textes de références, définitions etc. <http://openaccess.inist.fr/>

² Aujourd'hui, arXiv est hébergé par l'Université Cornell à Ithaca, NY <http://arxiv.org/>

³ La plateforme HAL <http://hal.archives-ouvertes.fr/>

⁴ Le site d'Érudit <http://www.erudit.org/>

⁵ Le site de Revues.org <http://www.revues.org/>

⁶ Le site de SciELO <http://www.scielo.org/>

⁷ Comme le modèle *OpenEdition Freemium* du Centre pour l'édition électronique ouverte (CLEO) qui gère Revues.org <http://www.openedition.org/8873?lang=fr>

l'offre commerciale des éditeurs, qui désormais font payer les auteurs (ou leurs institutions) pour la libre diffusion de leurs publications.

Stevan Harnad, l'un des protagonistes de ce mouvement, définissait ces deux approches complémentaires comme « green road » (voie verte) et « gold road » (voie dorée) (Harnad et al. 2008). La voie verte désigne l'auto-archivage des publications dans les plateformes d'auto-archivage aussi appelées « archives ouvertes⁸ » (arXiv, HAL etc.), tandis que la voie dorée correspond à la publication d'articles dans les revues en libre accès (Érudit, Revues.org, etc.).

Depuis la déclaration de Berlin sur le libre accès de 2003⁹ qui définissait le libre accès comme « une source universelle de la connaissance humaine et du patrimoine culturel ayant recueilli l'approbation de la communauté scientifique », le nombre d'archives ouvertes a dépassé les 3 000 sites¹⁰, et on compte plus de 8 000 revues en libre accès¹¹ sans qu'il soit possible de savoir combien d'articles, communications, thèses, etc. sont désormais librement accessibles. Un rapport récent de la Commission Européenne a estimé le pourcentage d'articles scientifiques en *open access* à 50% (Archambault et al. 2013)¹² mais ce chiffre varie en fonction des disciplines et pays et il est probablement surestimé. Cependant il est certain qu'une partie croissante de la communication scientifique est devenu « bien public » avec une diffusion plus rapide et plus vaste qu'auparavant (Willinsky 2005).

Loin d'être uniforme ou coordonné, l'ensemble de ce mouvement est riche et varié. Du point de vue technique et économique, la diversité est la règle, avec beaucoup d'initiatives locales de la part des universités et structures de recherche, mais aussi un investissement toujours plus grand de la part des éditeurs scientifiques classiques. Notre propre analyse des archives ouvertes révèle une grande variété de contenu qui recouvre en fait une réalité hétérogène (Schöpfel & Prost 2010 et 2013). On y trouve des archives thématiques pour un domaine ou une discipline spécifique (comme ArchiveSIC), des archives à vocation nationale comme par exemple HAL, des archives institutionnelles pour les publications des chercheurs d'un établissement (comme Archimer de l'IFREMER), puis d'autres sites à caractère scientifique qui n'appartiennent à

⁸ Un mot sur la terminologie. Ce terme est utilisé dans le sens des termes anglais *archive* ou *repository*. Il « désigne un réservoir où sont déposées des données issues de la recherche scientifique et de l'enseignement et dont l'accès se veut ouvert c'est-à-dire sans barrière. Cette ouverture est rendue possible par l'utilisation de protocoles communs qui facilitent l'accessibilité de contenus provenant de plusieurs entrepôts maintenus par différents fournisseurs de données. Synonymes : archive OAI ; archives en libre accès ; archives électroniques ouvertes. » Source : <http://openaccess.inist.fr/spip.php?page=glossaire> Il ne faut surtout pas faire l'amalgame avec le concept des archives traditionnelles, au sens patrimoine, préservation pérenne, etc.

⁹ <http://openaccess.inist.fr/?Declaration-de-Berlin-sur-le-Libre>

¹⁰ Source : les répertoires internationaux d'archives ouvertes *Directory of Open Access Repositories* OpenDOAR <http://www.opendoar.org/> et *Registry of Open Access Repositories* ROAR <http://roar.eprints.org/>

¹¹ Source : DOAJ <http://www.doaj.org/>

¹² Source : <http://openaccess.inist.fr/?Pres-de-50-des-publications-sont>

aucune de ces catégories (comme par exemple Symposcience ou les Bibliothèques Virtuelles Humanistes) (Armbruster & Romary 2009), comprenant des travaux de recherche, des données, des cours, des œuvres artistiques ou d'autres contenus (Hitchcock & Tarrant 2011). Parallèlement, on assiste à une professionnalisation et une centralisation de l'offre de services¹³, dont l'objectif est de réaliser des économies d'échelle, de simplifier l'accès à l'information et de développer une offre de service de qualité.

Il est donc temps de dresser un bilan et d'évaluer les résultats de l'investissement public en termes d'influence, de visibilité et d'usage. Néanmoins, il manque actuellement d'outils théoriques et pratiques pour procéder à une évaluation en bonne et due forme. Surtout, un référentiel fait défaut, un ensemble structuré de recommandations ou de bonnes pratiques, une *check-list* donnant des critères validés et structurés, utile aussi bien pour la création et gestion que pour l'audit d'un site.¹⁴

Certes, il existe le protocole OAI-PMH¹⁵, le Dublin Core¹⁶, les critères de sélection du répertoire OpenDOAR¹⁷ et d'autres projets et études, en particulier pour les archives institutionnelles (Lynch 2003, Vernooy-Gerritsen et al. 2010, Carr et al. 2008). Leurs critères de sélection s'appuient en règle générale sur cinq aspects : le lien avec l'établissement, le contenu, la garantie du libre accès, l'interopérabilité et les services à valeur ajoutée (évaluation, bibliographie, alertes, etc.).

Ces listes restent toutefois assez abstraites, parfois superficielles et non-structurées. Or, il existe un référentiel construit et organisé comme une norme, qui pourrait s'avérer utile, un peu comme un point de repère : il s'agit du certificat *Dokumenten- und Publikationsservice* de l'Initiative Allemande pour l'Information en Réseau DINI¹⁸. Voici une description de ce certificat dans son contexte, suivie de quelques pistes de réflexion.

¹³ Cf. en France, le protocole d'accord signé en 2006 et renouvelé en 2013 entre les universités et de nombreux organismes de recherche afin que la plateforme HAL devienne une plateforme commune à forte visibilité internationale : <http://www.bibliothequescientifiquenumerique.fr/?BSN4-Les-archives-ouvertes,52#nb1>

¹⁴ L'outil *Digital Repository Audit Method Based on Risk Assessment* (DRAMBORA) propose une solution mais n'est pas spécifique aux archives ouvertes (Pejsova & Vaska, 2011).

¹⁵ « Le protocole OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*) est un protocole qui définit les conditions du transfert de métadonnées d'une archive ouverte, produite par un fournisseur de données, vers le serveur d'un fournisseur de services ». Source : <http://openaccess.inist.fr/spip.php?page=glossaire>

¹⁶ Un ensemble de 15 éléments descriptifs (métadonnées) d'une ressource numérique cf. <http://dublincore.org/>

¹⁷ Voir sur le site déjà mentionné <http://www.opendoar.org/suggestionschart.html>

¹⁸ *Deutsche Initiative für Netzwerkinformation* <http://www.dini.de/>

Le contexte et l'histoire

L'association allemande DINI a été créée en 2001 par des bibliothèques, des centres de calculs et des services audiovisuels et multimédia universitaires. Son but : améliorer et intégrer l'infrastructure informationnelle de l'enseignement supérieur et de la recherche. DINI est un forum de discussion et de partage d'expériences sur le développement et l'environnement des technologies de l'information et de la communication. Parallèlement, DINI fait la promotion de solutions et de bonnes pratiques dans le domaine de l'infrastructure de l'information scientifique.

DINI s'appuie surtout sur l'élaboration de normes et de recommandations par différents groupes de travail thématiques. L'un des groupes de travail est consacré à l'édition numérique. Ce groupe existe depuis la création de DINI et est à l'origine des trois versions du certificat des archives ouvertes, mais aussi d'une terminologie pour les différents types de documents et publications ainsi que de recommandations pour les statistiques d'utilisation des archives ouvertes. La contribution aux projets de DINI est volontaire, honorifique et non rémunérée ; les membres des comités et des groupes de travail sont souvent employés par les 140 établissements membres.

Le point de départ du certificat fut, en 2002, un rapport émettant des recommandations pour l'édition numérique dans l'enseignement supérieur¹⁹. Ce document désignait les éléments importants à prendre en compte pour la création d'infrastructures durables dans le domaine de l'édition numérique scientifique. L'idée était d'aider les universités et autres établissements scientifiques à mettre en place leurs propres archives ouvertes et plateformes de revues, et de contribuer à un ensemble cohérent, compatible et interopérable. Il s'agit surtout d'éléments d'ordre organisationnel et juridique, mais aussi technique. Le rapport a été rédigé par le groupe de travail consacré à l'édition numérique, qui fut l'un des premiers groupes créés par le DINI et qui est toujours actif. De ce rapport est née l'idée d'un catalogue de critères pour évaluer la qualité du serveur d'une archive ouverte. Ce catalogue a pris la forme d'un certificat dont la première version à sept critères a été publiée en 2003 sous le nom *DINI-Zertifikat Dokumenten- und Publikationsserver 2004*²⁰. Deux nouvelles versions en ont élargi et précisé les critères : le certificat de 2007 a mis l'accent sur l'ensemble du service ; celui de 2010 a modifié en profondeur la structure des critères, en ajoutant une numérotation univoque des conditions afin de faciliter le référencement et de permettre une utilisation sous forme d'une liste de vérification (*check-list*)²¹. Du point de vue du contenu, la dernière version a surtout changé les aspects juridiques, la sécurité de l'information et l'interface

¹⁹ DINI (2002) *Elektronisches Publizieren an Hochschulen* <http://nbn-resolving.de/urn:nbn:de:kobv:11-10045992>

²⁰ *DINI-Zertifikat Dokumenten- und Publikationsserver* <http://nbn-resolving.de/urn:nbn:de:kobv:11-10046067>

²¹ *DINI-Zertifikat Dokumenten- und Publikationsservice 2010* <http://nbn-resolving.de/urn:nbn:de:kobv:11-100182794>

OAI-PMH. Elle a également commencé à intégrer les revues en libre accès (voie dorée) et les résultats scientifiques (*datasets*) même si les documents textuels et les archives institutionnelles et thématiques restent toujours au cœur du certificat.

Le groupe de travail

Le groupe de travail sur l'édition numérique est composé de représentants des bibliothèques et des centres de calculs, issus pour la plupart d'établissements ayant des archives ouvertes. Le résultat de leur travail reflète ainsi les connaissances théoriques et beaucoup d'expérience pratique, ce qui justifie son caractère de « guide de bonnes pratiques ».

Sur demande d'un établissement responsable d'une archive ouverte, deux membres du groupe de travail prennent en charge la certification d'un site, c'est-à-dire qu'ils évaluent à quel degré il remplit les critères du certificat. L'évaluation se fait dans l'esprit d'une revue par les pairs (*peer review*) dans la mesure où experts et demandeurs appartiennent à la même communauté, non anonyme, sous forme d'un processus coopératif dans lequel les deux partis cherchent ensemble comment le site peut remplir les conditions du certificat, une étape qui dure deux mois en moyenne.

Le retour d'expérience de ce partenariat au quotidien, de la part des experts évaluateurs mais aussi de la part des professionnels candidats à la certification - contribue à faire évoluer le contenu du certificat. Lors de la préparation d'une nouvelle version, ce retour d'expérience est organisé d'une manière explicite et systématique, suscitant de nombreux commentaires, suggestions et remarques. De même, à l'occasion de la nouvelle version en 2013, DINI a procédé à une analyse critique des résultats de la certification, afin d'éliminer des recommandations jamais suivies ou, au contraire, mettre à jour d'autres conditions « trop faciles ». Quelques exemples :

- Le groupe de travail s'informe régulièrement des changements en matière de normes techniques, notamment dans les domaines de la sécurité de l'information et de l'archivage pérenne (critères 5 et 8), et adapte le certificat en fonction de ces changements.
- La question de l'identifiant pérenne a suscité beaucoup de débat controversé, et la nouvelle version 2014 abandonnera la recommandation actuelle qui demande un identifiant sous forme d'une adresse URL opérable (critère 5).
- Suite aux remarques de la part des opérateurs et services qui travaillent régulièrement avec les interfaces OAI (critère 6), DINI fait évoluer ses recommandations afin de tenir compte de la réalité des flux de données (moissonnage).
- Les recommandations concernant les statistiques d'utilisation (critère 7) évoluent en fonction des résultats et développements d'un projet de recherche sur les statistiques des archives ouvertes.

- Dernier exemple, suite aux retours des opérateurs et professionnels le groupe de travail a ajouté les services liés aux revues en libre accès au périmètre de la dernière version.

L'objectif

L'objectif du certificat DINI est de décrire les lignes directrices pour l'exploitation des archives ouvertes scientifiques. Faisant partie de l'infrastructure de publication des établissements scientifiques, ces plateformes doivent répondre à d'autres exigences techniques et organisationnelles que celles d'un simple site Web personnel. Cependant, la question de la transposition des exigences et des normes du système établi de la publication scientifique (telles que la fiabilité, la durabilité, la « citabilité » et l'interopérabilité globale) dans ce nouvel environnement numérique est encore ouverte à débat.

Le certificat DINI contribue à ce débat d'une manière importante. S'il sert principalement de repère pour la certification des archives ouvertes des universités et des sociétés savantes, il joue aussi le rôle d'un référentiel pour la création ou l'amélioration d'une telle plateforme. Par ailleurs, il peut servir d'argument de poids pour une demande budgétaire puisqu'il illustre l'investissement nécessaire pour atteindre un niveau de qualité exigé et acceptable.

Le certificat DINI contient une description précise des critères de conformité avec les normes et développements nationaux et internationaux dans le domaine des systèmes d'information et sites du libre accès et du Web. Pour les services qui remplissent les conditions nécessaires, ce label de qualité augmentera autant leur visibilité au sein de leur établissement ou de leur discipline que sur le plan international.

La certification DINI est censée garantir l'interopérabilité des services, c'est-à-dire leur capacité à fonctionner ensemble et à contribuer à des services intégrés, comme par exemple des moteurs de recherche ou répertoires de sites qui permettent d'interroger le contenu de plusieurs centaines d'archives ouvertes d'une manière simultanée, comme BASE²² ou OAlster²³ qui couvrent plusieurs dizaines de millions de documents en libre accès. Un minimum de cohérence et de compatibilité au niveau des formats et données est indispensable pour faciliter l'accès à l'information. C'est l'objectif du certificat.

La procédure de la certification n'est pas limitée aux établissements de l'enseignement supérieur et de recherche. Des éditeurs *open access*, des

²² Le métamoteur de recherche BASE (Bielefeld Academic Search Engine), mis en place par l'université de Bielefeld en Allemagne, recense plus de 50 millions de documents en libre accès (septembre 2013) <http://www.base-search.net/>

²³ La base de données OAlster avec plus de 23 millions de références (août 2013) est aujourd'hui gratuitement accessible via l'offre de service d'OCLC <http://www.oclc.org/oaister.en.html?urlm=168646>

agrégateurs et des centres de calcul sont également invités à contribuer à la communication scientifique de demain et à demander la certification DINI.

Par le biais de la certification, DINI propose un contrôle transparent de qualité des archives ouvertes. Le certificat réunit ainsi les conditions indispensables à une communication scientifique moderne.

La structure

Au cœur du certificat se trouve un catalogue de critères comportant les conditions techniques, organisationnelles et juridiques qu'une archive ouverte devrait remplir. Le certificat part du principe que les normes en vigueur pour la publication scientifique s'appliquent aussi au contexte numérique, comme la crédibilité, la fiabilité, la durabilité, l'accessibilité, la pérennité et la « citabilité » des documents. Mais le certificat va plus loin dans la mesure où il ajoute une prestation d'assistance et de conseil pour les auteurs et les éditeurs, mettant ainsi de l'avant l'importance de l'offre de service de tels systèmes de publication.

Le catalogue contient huit critères regroupant les conditions minimales nécessaires pour l'obtention du certificat. Chaque condition est clairement identifiée avec un intitulé, une numérotation et une formulation qui permettent une application sous forme de *check-list* (question/réponse de type oui/non).

Les conditions minimales sont suivies par des recommandations qui n'ont pas de caractère obligatoire mais qui indiquent toutefois des tendances et des suggestions importantes susceptibles de devenir des conditions obligatoires par la suite.

Toutes les conditions et les recommandations sont expliquées et illustrées à l'aide d'exemples et de modèles concrets.

À la version 2010 du certificat correspond un questionnaire en ligne que l'établissement demandeur doit remplir pour lancer la démarche de certification.

Les critères

Comme évoqué plus haut, le certificat DINI regroupe huit critères d'évaluation. Voici sous forme de tableau un premier aperçu de ces critères, avec le nombre de conditions et, le cas échéant, les sous-domaines.

Numéro	Critères du certificat DINI	Nombre de conditions
1	Visibilité du site	4
2	Charte	8
3	Aides aux auteurs et éditeurs	4
4	Aspects juridiques	10
5	Sécurité de l'information	13

5.1	<i>Système technique</i>	8
5.2	<i>Traitement des documents</i>	5
6	Indexation et interfaces	7
6.1	<i>Indexation</i>	5
6.2	<i>Interfaces de métadonnées</i>	2
7	Statistiques d'utilisation	4
8	Archivage pérenne	2

Les critères sont plus ou moins exhaustifs et détaillés. Certains, comme pour la visibilité du site, l'aide aux auteurs et éditeurs ou l'archivage pérenne, sont relativement simples. D'autres, notamment pour les aspects juridiques et techniques (sécurité, interfaces) sont plus élaborés. Voici les grandes lignes du contenu de ces critères. L'annexe reprend cette liste, avec la numérotation et la formulation précise de chaque condition minimale pour la certification.

Visibilité du site

Le critère de la visibilité du site fait partie du certificat depuis 2007 pour deux raisons : souligner l'importance stratégique d'une archive ouverte pour l'image et l'influence d'un établissement scientifique et faire connaître ce service auprès des chercheurs de l'établissement. Le critère demande à ce que le service soit bien visible et facilement accessible à partir des pages centrales de l'établissement ; il faut aussi que le site et son interface OAI (pour le moissonnage automatique) soient répertoriés dans les annuaires et les services habituels (critère 6), notamment dans la liste du DINI, le site OpenDOAR et le registre officiel de l'initiative OAI.

Charte

La transparence des services est primordiale pour la crédibilité d'un site de publication. La notion de transparence est présente dans tout le certificat mais s'exprime tout particulièrement dans l'existence et la communication d'une charte. Cette dernière précise les obligations des auteurs, des éditeurs et des hébergeurs et détermine la politique éditoriale, c'est-à-dire les documents et les contenus acceptés ainsi que le public cible. De plus, la charte doit informer sur la durée garantie de l'accès au service, sur la pérennité de l'archivage et sur le fonctionnement technique. L'actuelle version du certificat demande également un engagement en faveur du libre accès mais n'exige pas que tous les documents soient en libre accès. Ceci étant, la majorité des documents doit être librement accessible dans le cas contraire, la charte doit expliquer quels documents ne le sont pas et pourquoi.

Aide aux auteurs et éditeurs

Une archive ouverte est tout d'abord une incitation aux auteurs et aux éditeurs à déposer et publier leurs travaux scientifiques sur une plateforme fiable. Le détail de cette proposition et l'assistance proposée aux utilisateurs doivent donc être

explicites. Cependant, le certificat reste relativement vague à cause de la diversité des sites : une archive pour les documents déjà publiés doit proposer un autre type d'assistance qu'une plateforme de revues. Dans tous les cas, le site doit fournir une adresse de correspondance (contact) pour répondre aux questions et aux problèmes. Dans d'autres cas, il suffira de renvoyer vers des sites tiers, par exemple SHERPA/RoMEO²⁴ pour les aspects juridiques d'un dépôt. Les auteurs doivent avoir la possibilité de déposer eux-mêmes leurs propres travaux, sauf si ce dépôt est assuré par un service central, par exemple la bibliothèque de l'établissement.

Aspects juridiques

L'hébergeur d'une archive ouverte a besoin de certains droits d'exploitation pour la mise à disposition et la dissémination d'un document. Le certificat demande une « licence de dépôt » pour la cession de ces droits au moment du dépôt. Il n'est pas nécessaire de signer un document imprimé ; il suffit d'accepter les termes de la licence en ligne. Les métadonnées de chaque document doivent renvoyer vers la licence pour indiquer quels droits ont été réellement cédés, surtout s'il existe plusieurs versions de la licence.

Sécurité de l'information

Si elle veut être fiable, une archive ouverte doit remplir les conditions élémentaires de la gestion d'un système d'information. La documentation joue ici un rôle primordial. La description et la documentation des choix et des paramétrages du système ainsi que des règles de gestion, de maintenance et de sécurité sont indispensables pour assurer un service stable et durable. Ces conditions concernent entre autres choses l'existence d'une approche globale de sécurité, la désignation d'un responsable de la sécurité, l'existence d'une procédure concept opérationnelle et de sauvegarde et d'un plan de continuité d'activité en cas d'incident. Par ailleurs, il faut définir l'accès aux différentes composantes du système technique et assurer un contrôle automatique et régulier de la disponibilité de ces composantes.

Un deuxième aspect de la sécurité concerne le traitement des documents afin d'en assurer l'intégrité et la « citabilité », par analogie aux publications imprimées. En particulier, les conditions précisent qu'un document déposé et publié ne doit plus être modifié mais peut être, le cas échéant, complété par une nouvelle version, et qu'il doit disposer d'un identifiant pérenne (*persistent identifier*). Pour assurer l'intégrité durable des publications, le certificat recommande aussi l'application d'une fonction de hachage sécurisé (*hash*).

Indexation et interfaces

Afin de faciliter l'accès à l'archive par le biais de moteurs de recherche et l'interopérabilité avec d'autres sites, le certificat recommande de la transparence

²⁴ *Publisher copyright policies & self-archiving* <http://www.sherpa.ac.uk/romeo/>

au niveau de l'indexation, de rendre l'indexation compatible avec le Dublin Core (simplifié) et d'autres terminologies normalisées, et de mettre les métadonnées à disposition via une interface OAI.

L'interface OAI est décrite dans l'annexe A du certificat. Elle doit être conforme au protocole OAI-PMH. Par ailleurs, l'annexe A définit plusieurs ensembles de données (*sets*) afin de garantir l'interopérabilité et le moissonnage des métadonnées. Elle décrit également la procédure à suivre pour les documents supprimés et le contrôle du flux des données. Elle spécifie également les champs obligatoires et la structure interne de certaines métadonnées, par exemple pour la date de publication et la langue du document.

Ces conditions sont compatibles avec les travaux et les recommandations du projet européen DRIVER.

Statistiques d'utilisation

Le certificat DINI prescrit d'appliquer une méthode normalisée pour mesurer l'utilisation des documents et d'en communiquer les résultats pour chaque document. L'objectif est de répondre adéquatement à la demande d'une évaluation quantitative de la diffusion et de l'influence de la production scientifique. Mettre à disposition les statistiques d'utilisation d'une manière transparente et compatible avec d'autres sites contribue au système d'évaluation des publications scientifiques et peut en outre créer un incitatif à publier sur cette plateforme.

Néanmoins, dans la mesure où ce domaine est encore peu développé, le certificat émet surtout des recommandations qui s'appuient sur les résultats du projet DINI OA-Statistik, en particulier sur la normalisation des indicateurs et sur la collecte et le traitement centralisés des journaux (*log files*).

Archivage pérenne

La garantie de pérennité contribue à la crédibilité d'une archive ouverte. Afin d'assurer qu'une publication scientifique puisse être vérifiée et citée, elle doit rester accessible à long terme dans sa forme et dans sa structure initiale. Pour un document numérique, cela a trois significations : conserver le format (c'est-à-dire l'exakte séquence des bits), le rendre accessible et en préserver l'intelligibilité, c'est-à-dire l'interprétation correcte des données binaires au sens de la représentation de l'information et de la mise à disposition d'autres fonctions²⁵.

Même s'il existe aujourd'hui de nombreux projets et initiatives dans ce domaine, la mise en œuvre de ces concepts et de ces technologies en est encore à ses débuts et encore peu d'établissements sont aptes à proposer une solution fiable

²⁵ Cf. l'explication plus détaillée sur le site du Centre Informatique National de l'Enseignement Supérieur (CINES) en France <http://www.cines.fr/spip.php?rubrique229>

pour l'archivage pérenne. Le certificat s'est donc abstenu d'imposer des conditions et a préféré suggérer une coopération avec des prestataires compétents²⁶. En attendant des solutions viables, le certificat se contente de demander à l'hébergeur la garantie d'une durée de conservation minimale de cinq ans²⁷ et l'absence de mesures techniques de protection (*digital rights management* ou DRM, protection par mot de passe, etc.).

Certification, exemples et annexes

Le certificat contient trois autres sections : la procédure de certification par le DINI ; des explications complémentaires et des exemples pour tous les critères ; trois annexes : les interfaces OAI du site (annexe A), un glossaire des termes techniques (annexe B) et une liste complète des auteurs ayant contribué à sa rédaction et à sa mise à jour (annexe C).

La procédure de certification et les explications complémentaires sont assez étroitement liées au contexte allemand du certificat. L'annexe A paraît toutefois indispensable à la création et à l'évaluation d'une archive ouverte dans la mesure où il s'agit de rendre un site conforme avec le protocole international OAI-PMH. De même, le glossaire (annexe B) est certainement d'une grande utilité pour la compréhension générale du certificat mais aussi, dans une version bilingue ou trilingue, pour la traduction et l'adaptation du texte dans d'autres langues.

Discussion

L'objet de certification

Initialement, le certificat DINI fut développé pour les archives ouvertes dans l'esprit de la voie verte²⁸, pour accompagner la mise en place des archives institutionnelles par les universités et organismes de recherche allemands. Par la suite, les experts du réseau DINI ont été de plus en plus sollicités pour évaluer également des plateformes de revues en libre accès (voie dorée). DINI a interprété ce développement comme une convergence partielle et a pris dès lors deux décisions : ouvrir le certificat aux nouveaux services d'information scientifique en libre accès mais écarter, au moins pour un temps, l'idée d'un certificat modulaire (ou à plusieurs options). Ce choix pose deux problèmes :

(a) Quel est l'objet du certificat? La première version parle d'archive ouverte ; la troisième utilise l'expression « service de documents et de publications » tout en restant centrée sur les archives institutionnelles. Ce terme n'a pas d'équivalent en français. Faut-il parler de plateforme documentaire ? de service de publication

²⁶ En Allemagne, il s'agit en premier lieu de la bibliothèque nationale qui développe actuellement un portefeuille de services (projet DP4lib). En France, on penserait plutôt au Centre Informatique National de l'enseignement supérieur (CINES) : <http://www.cines.fr/>

²⁷ Ce qui est peu, comparé à la garantie d'une préservation de 30 ans par le CINES en France.

²⁸ Rappelons que la voie verte ou « green road » désigne l'auto-archivage des articles dans les archives ouvertes, par les chercheurs.

en libre accès? de site ou serveur OA *open access* ? Dans cet article, nous avons privilégié le terme d'archive ouverte mais nous sommes conscients du fait qu'il est réducteur.

(b) À cause du grand écart entre les voies verte et dorée, la troisième version du certificat reste parfois trop abstraite : certains critères nécessitent une adaptation au terrain, à la finalité, au public et à la communauté scientifique du site en question.

En ceci, le certificat reflète sans doute aussi une période de transition entre le système traditionnel de la communication scientifique, basé pour l'essentiel sur des revues et monographies payantes, et le nouveau paysage de l'information scientifique profondément marqué du libre accès en constante évolution mais sans modèle prépondérant.

L'influence du certificat

L'accueil du certificat par les professionnels a été favorable. À ce jour, environ 45 archives ouvertes ont reçu leur certification, soit le quart des sites répertoriés par le DINI²⁹. Certains ont renouvelé la certification pour chaque nouvelle version du certificat, d'autres non. Presque toutes les démarches aboutissent. En effet, le DINI s'efforce d'aider les établissements à adapter leurs sites en conformité avec les conditions du certificat plutôt que de décerner un label sélectif.

Cet objectif implique une activité soutenue d'information, de communication et de promotion par le biais d'ateliers et de tutoriels, de conférences, de sessions par affiche et de publications, ainsi qu'un contact direct avec les institutions et professionnels et l'accompagnement d'initiatives locales. Cette activité est partiellement financée par l'Agence nationale de recherche allemande (DFG)³⁰.

Bien que le nombre réel de services certifiés reste encore relativement modeste, l'attention dont bénéficie le catalogue des critères est tout à fait remarquable, tant au niveau national (allemand) qu'au niveau international. À preuve, le certificat a été traduit en anglais, en français et en espagnol, et ces trois versions sont accessibles sur le serveur de l'Université Humboldt à Berlin³¹.

Le label qualité DINI renforce sans conteste la crédibilité et la visibilité des sites et facilite leur mise en réseau. L'élément névralgique demeure néanmoins le retour sur l'investissement : l'effort nécessaire est substantiel et les retombées ne sont pas immédiates. Mais même si le contexte allemand n'est pas transposable dans d'autres pays, il n'en reste pas moins que ce certificat est à ce jour le seul document cohérent et construit à proposer un catalogue de critère pour évaluer et faire évoluer les nouveaux sites du libre accès.

²⁹ Liste des sites répertoriés par DINI <http://www.dini.de/dini-zertifikat/liste-der-repositorien/>

³⁰ Deutsche Forschungsgemeinschaft : <http://www.dfg.de/>

³¹ Version française : <http://edoc.hu-berlin.de/series/dini-schriften/2010-3-fr/PDF/dini-zertifikat-2010-3-fr.pdf>

Normalisation et particularités

Le certificat aspire au statut d'outil de référence en matière d'archives ouvertes. Il est plus qu'un simple répertoire de « bonnes pratiques ». Grâce à son caractère de (quasi) norme, son mode d'élaboration, sa présentation et son application le rapprochent très clairement des projets comme COUNTER³² qui, il y a quelques années, a établi d'une manière collective et consensuelle des recommandations fortes dans le domaine des statistiques d'utilisation des ressources en ligne (revues, livres, bases de données, ...), en imposant des procédures et une terminologie acceptées par les principaux acteurs (éditeurs, agences d'abonnements, agrégateurs, consortia de bibliothèques, centres d'information scientifique, etc.).

Cela étant, à la différence de COUNTER, le certificat distingue trois niveaux de recommandations dont seul le premier est obligatoire, tandis que les autres restent facultatifs et indicatifs. De même, là où COUNTER et d'autres initiatives de normalisation se placent (malgré un ancrage fort dans les pays anglophones) essentiellement au niveau international, le certificat DINI contient des références fortes à l'environnement allemand. Par exemple, la description des conditions et des recommandations renvoie plus de vingt fois à des lois, des établissements, des formats, des outils et des sites allemands. Dans la traduction (cf. annexe), nous avons marqué ces renvois par un astérisque car la transposition dans un autre pays doit tenir compte du contexte national spécifique.

Dernière remarque : à quelques exceptions près, les archives ouvertes et les plateformes de revues font partie du système d'information d'un établissement scientifique et sont donc, d'une manière ou d'une autre, connectées à un système de recherche. Or, ces systèmes évoluent vers d'avantage d'intégration, d'interopérabilité et de normalisation, par exemple autour du format européen CERIF³³. CERIF sert de format pivot pour un nombre croissant de systèmes d'information de recherche (*current research information systems* ou CRIS) qui facilitent la gestion et le pilotage des projets et établissements scientifiques. Vu sous cet angle, il est tout à fait possible que les archives ouvertes intègrent à leur tour l'architecture des systèmes CRIS, qu'elles adoptent le format CERIF (ou un équivalent) au détriment des métadonnées du Dublin Core ou qu'elles communiquent avec d'autres composantes du système (finances, ressources humaines, équipements, projets, etc.) via CERIF comme format d'échange du logiciel médiateur (*middleware*). Le rapprochement entre les communautés des archives ouvertes et des systèmes d'information de recherche est en cours (cf.

³² Projet COUNTER (*Counting Online Usage of Networked Electronic Resources*)
<http://www.projectcounter.org/>

³³ *Common European Research Information Format*, maintenu et développé par EuroCRIS
<http://www.eurocris.org/>

les communications des conférences internationales CRIS³⁴), et l'avenir dira comment et à partir de quels données et formats cette intégration se fera.

Conclusion

Après une introduction au mouvement vers le libre accès à l'information scientifique, notre article a présenté le certificat du réseau allemand DINI comme une approche intéressante pour évaluer les archives ouvertes qui représentent l'une des voies majeures de ce mouvement. Nous avons placé ce certificat dans le contexte du réseau DINI, et nous avons essayé d'en dégager les grandes lignes.

Les enjeux sont de taille. L'évolution rapide des technologies de l'information et des modes de la communication scientifique oblige le certificat à s'adapter continuellement, y compris à élargir son périmètre initial afin d'intégrer de nouveaux sites en libre accès. L'application de ce certificat demande un investissement non négligeable de la part des institutions, et cette contrainte freine quelque peu le succès de cet outil. De même, une telle approche implique également un travail d'adaptation au contexte national ou régional ; le certificat a été développé en Allemagne. Une application en France ou au Canada nécessite un certain travail sur le contenu des critères, conditions et recommandations. Ce travail reste à faire.

Nous avons parlé plus tôt de la grande diversité des archives ouvertes. Une approche telle que le certificat DINI n'entre pas en contradiction avec cette richesse et cette variété. Il ne s'agit pas d'uniformiser le paysage des archives ouvertes mais de créer un cadre référentiel qui permet d'évaluer les différents projets et qui, en favorisant l'interopérabilité, œuvre en faveur d'une interconnexion et d'une mise en réseau progressive des sites. L'intérêt est double : contribuer à une évaluation des résultats d'une politique publique en faveur du libre accès, et favoriser l'émergence d'une offre de service de qualité.

Quel est l'avenir du certificat ? DINI prépare une nouvelle version. La révision s'appuie sur l'expérience de la certification et précisera certains critères et conditions; par exemple, le signalement d'un site dans le répertoire OpenDOAR restera une recommandation mais ne sera plus obligatoire. Parallèlement, la nouvelle version appelée « Open-Access-Repositorien und – Publikationsdienste³⁵ » ira plus loin dans les spécifications destinées aux plateformes de revues en libre accès, par exemple pour la distinction entre contact technique et contact éditorial. Elle rentrera en vigueur au courant de 2014.

Fondamentalement, fonction et objectif du certificat resteront les mêmes : faciliter l'évaluation de ces nouveaux services, contribuer à leur interopérabilité et

³⁴ Voici le site de la dernière conférence CRIS 2012 à Prague, avec le programme et les liens vers les communications et présentations <http://www.cris2012.org/>

³⁵ En français : « archives ouvertes et services d'édition en libre accès ».

accompagner le développement du libre accès dans l'univers de l'information scientifique. On peut avoir des réticences à l'idée d'une normalisation ou d'une procédure formalisée ; néanmoins, même en tant que « bonnes pratiques » les recommandations du certificat DINI restent extrêmement utiles pour tous les acteurs en charge de tels projets.

Remerciements à Ulrich Herb (Université du Saarland) et Frank Scholze (Karlsruhe Institute of Technology) pour leurs conseils lors de la rédaction de l'article.

Tous les sites et les documents de référence ont été consultés le 10 mars 2014.

Bibliographie

Archambault, Eric, Didier Amyot, Philippe Deschamps, Aurore Nicol, Lise Rebout, et Guillaume Roberge. "Proportion of Open Access Peer-Reviewed Papers at the European and World Levels — 2004-2011." Rapport, Commission Européenne DG Recherche & Innovation, août 2013.
<http://www.science-metrix.com/pdf/SM_EC_OA_Availability_2004-2011.pdf>

Armbruster, Chris et Laurent Romary. "Comparing Repository Types: Challenges and Barriers for Subject-Based Repositories, Research Repositories, National Repository Systems and Institutional Repositories in Serving Scholarly Communication." *Social Science Research Network Working Paper Series*, novembre 2009.
<http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1506905>

Carr, Les, Wendy White, Susan Miles, et Bill Mortimer. "Institutional Repository Checklist for Serving Institutional Management." *Third International Conference on Open Repositories 2008, 1-4 April 2008, Southampton, United Kingdom*. 2008.
<<http://eprints.soton.ac.uk/23929/>>

DINI. Elektronisches Publizieren an Hochschulen - Empfehlungen. Rapport, Deutsche Initiative für NetzwerkInformation, 2002.
<<http://nbn-resolving.de/urn:nbn:de:kobv:11-10045992>>

Harnad, Stevan, Tim Brody, François Vallières, Les Carr, Steve Hitchcock, Yves Gingras, Charles Oppenheim, Chawki Hajjem, et Eberhard R. Hilf. "The Access/Impact Problem and the Green and Gold Roads to Open Access: An Update." *Serials Review* 34 (2008): 36-40.
<<http://depts.washington.edu/uwbri/PDF%20Files/access-impact%20problem%20and%20the%20green%20and%20gold%20roads%20to%20open%20access-an%20update.pdf>>

- Hitchcock, Steve, et David Tarrant. "Characterising and Preserving Digital Repositories: File Format Profiles." *Ariadne* 66 (2011): 30 janvier 2011.
<<http://www.ariadne.ac.uk/issue66/hitchcock-tarrant>>
- Lynch, Clifford A. Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age. Rapport 226, ARL Association of Research Libraries, février 2003.
<<http://www.arl.org/resources/pubs/br/br226/br226ir.shtml>>
- Mahé, Annaig. "Bibliothèques et archives ouvertes." *Bulletin des Bibliothèques de France* 56 (2011): 14-18.
<<http://bbf.enssib.fr/consulter/bbf-2011-01-0014-003>>
- Pejšová, Petra, et Marcus Vaska. "Audit DRAMBORA for trustworthy repositories: A Study Dealing with the Digital Repository of GL." *Thirteenth International Conference on Grey Literature: The Grey Circuit. From Social Networking to Wealth Creation. Washington, DC, 5-6 December 2011*.
<<http://www.opengrey.eu/item/display/10068/700161>>
- Schöpfel, Joachim, et Hélène Prost. "Développement et Usage des Archives Ouvertes en France. Rapport. 1e partie : Développement." Rapport final, Université Charles-de-Gaulle Lille 3, 2010.
<http://archivesic.ccsd.cnrs.fr/sic_00497389/fr/>
- Schöpfel, Joachim, et Hélène Prost. "Archives institutionnelles : Observations sur un nouveau mode d'information scientifique." *Revue Canadienne des Sciences de l'Information et de Bibliothéconomie* 37 (2013): 122-136.
<doi:10.1353/ils.2013.0008>
- Suber, Peter. *Open access*. Cambridge Mass.: MIT Press, 2012.
<<http://mitpress.mit.edu/books/open-access>>
- Vernooy-Gerritsen, Marjan, Gera Pronk, et Maurits van der Graaf. "Three Perspectives on the Evolving Infrastructure of Institutional Research Repositories in Europe." *Ariadne* 59 (2009): 30 avril 2009.
<http://depot.knaw.nl/5660/1/Ariadne_issue59_vernooy-gerritsen-et-al_.pdf>

Annexe : Les critères sous forme d'une liste de vérification (check-list)

Voici, de manière succincte, les critères (c'est-à-dire les conditions minimales) du certificat DINI. Les explications et les recommandations sont disponibles sur le wiki du projet ainsi que sur le site du DINI. Les astérisques (*) marquent les critères liés au contexte allemand, critères à adapter dans un autre contexte national

1. Visibilité du site

M.1-1 Le service est accessible à partir d'une page Web.

M.1-2 Le lien vers la page d'accueil du service est accessible à partir du site institutionnel de l'établissement (ou des établissements) responsable(s) du site.

M.1-3 Le service est enregistré auprès du DINI dans la liste des plateformes de publication.*

M.1-4 Le service est enregistré auprès du répertoire Directory of Open Access Repositories (OpenDOAR).

2. Charte

M.2-1 L'établissement dispose d'une charte qui décrit la politique du service.

La charte contient les points suivants :

M.2-2 Une définition des droits et obligations de l'établissement.

M.2-3 Une définition des droits et obligations des auteurs et des éditeurs qui utilisent le service dans un but de publication.

M.2-4 Une description des documents publiés sur le site (typologie, qualité technique, contenu).

M.2-5 Une spécification de la durée minimale pour la mise à disposition des documents publiés sur la plateforme, avec la garantie correspondante.

M.2-6 Une déclaration concernant la conservation pérenne des documents.

M.2-7 Une explication du fonctionnement technique de l'archive institutionnelle.

M.2-8 Une déclaration solennelle concernant le libre accès (*open access*).

3. Aides aux auteurs et éditeurs

M.3-1 Il existe un service d'aide et d'assistance accessible à partir du site Web.

M.3-2 Les auteurs peuvent déposer leurs documents de façon autonome, par l'intermédiaire d'un formulaire Web ou par un autre biais.

M.3-3 Le site contient l'information pertinente sur tous les aspects techniques du dépôt. A défaut de quoi, si cette information se trouve déjà sur un autre site, il suffit de proposer un lien vers cette autre page.

M.3-4 Le site contient des liens vers des pages d'information juridique, notamment dans le cas du dépôt d'un document publié ailleurs (*post-print*).

4. Aspects juridiques

M.4-1 La relation juridique entre les auteurs ou les éditeurs (ayant droits) et l'établissement gestionnaire du site est régie par un contrat formel (cession de droits).

M.4-2 L'établissement diffuse sa licence de dépôt en ligne en version allemande.*

M.4-3 L'ayant droit cède à l'établissement le droit de stockage électronique, en particulier dans des bases de données, le droit de mise à disposition, le droit de diffusion publique et d'envoi et, dans cette optique, le droit d'archivage et de reproduction.

M.4-4 L'ayant droit cède à l'établissement le droit de signalement et de transfert à des tiers, par exemple dans le cadre d'une mission de collecte nationale, notamment dans le but de l'archivage pérenne.

M.4-5 L'ayant droit cède à l'établissement le droit de copie et de conversion vers d'autres formats numériques ou physiques pour assurer l'archivage pérenne, avec maintien de l'intégrité de l'œuvre.

M.4-6 L'ayant droit garantit à l'établissement que le dépôt du document (ou de parties du document, par exemple des images, etc.) ne porte pas atteinte aux droits d'une tierce partie (co-auteur, éditeur, établissement de financement, etc.).

M.4-7 L'ayant droit s'engage à informer l'établissement immédiatement en cas de litige ou d'autres problèmes juridiques réels ou supposés.

M.4-8 L'ayant droit dégage l'établissement de toute revendication éventuelle de la part d'une tierce partie.

M.4-9 Le site Web de l'archive contient un impressum en conformité avec la loi allemande sur les médias électroniques (TMG) et, le cas échéant, d'autres lois pertinentes.*

M.4-10 L'établissement ajoute le statut légal du document aux métadonnées.

5. Sécurité de l'information

5.1. Système technique

M.5-1 Le système d'information du site est intégré dans le dispositif de sécurité de l'établissement.

M.5-2 Un chargé de sécurité est nommé pour le système d'information.

M.5-3 Il existe des spécifications fonctionnelles avec un plan de maintenance du système d'information.

M.5-4 Un document écrit fournit le détail des spécificités du système d'information avec toutes les composantes nécessaires à son bon fonctionnement.

M.5-5 Il existe des règles écrites sur l'accès au système d'information et à ses composantes.

M.5-6 Tous les documents et données sont régulièrement sauvegardés (*backup*).

M.5-7 Il existe une procédure de gestion d'incident.

M.5-8 La disponibilité des serveurs du site est régulièrement contrôlée par un logiciel de surveillance autonome.

5.2. *Traitement des documents*

M.5-9 Les documents déposés sur le site ne sont plus modifiés par la suite.

M.5-10 Un identifiant pérenne est assigné à chaque document et à chaque nouvelle version.

M.5-11 Les identifiants pérennes sont utilisés dans les métadonnées exportées et sur le site comme identifiant primaire sous forme d'une URL opérationnelle.

M.5-12 La suppression d'un document reste exceptionnelle et doit être justifiée par écrit.

M.5-13 Lors de la mise en ligne d'un document déposé, le transfert de données se fait via le protocole SSL et à partir d'un certificat de confiance.

6. **Indexation et interfaces**

6.1. *Indexation*

M.6-1 Les règles d'indexation sont explicites et publiées sur le site.

M.6-2 Chaque document est indexé selon des outils et méthodes du *Dublin Core Element Set*.

M.6-3 Chaque document est indexé en fonction de son domaine (sujet ou thématique), soit avec une terminologie libre, soit à l'aide d'un vocabulaire contrôlé.

M.6-4 Chaque document est indexé selon la classification Dewey (DDC), au moins au niveau des groupes thématiques de la Bibliographie Nationale Allemande.*

M.6-5 Chaque document est indexé en fonction de son type suivant les recommandations du DINI pour une terminologie commune³⁶.*

6.2. *Interfaces de métadonnées*

M.6-6 Une interface Web donne accès à tous les documents et métadonnées.

M.6-7 Il existe une interface OAI compatible avec le protocole OAI-PMH 2.0 et avec les recommandations OAI du DINI.*

7. **Statistiques d'utilisation**

M.7-1 L'archive ouverte collecte des statistiques d'utilisation du site d'une façon fiable et en conformité avec la loi.

M.7-2 Les fichiers de connexion au site (*logs*) sont rendus anonymes pour un stockage à plus long terme.

M.7-3 L'accès automatique au site, aux documents ou aux données est filtré (robots).

M.7-4 Une documentation publique indique les critères et normes appliqués lors de la création des statistiques d'utilisation.

³⁶ <http://edoc.hu-berlin.de/series/dini-schriften/12/PDF/12.pdf>

8. Archivage pérenne

M.8-1 Le site précise la durée minimale de l'archivage et de la mise à disposition des documents et métadonnées publiés sur ce site. Cette durée est d'au moins cinq ans.

M.8-2 Les fichiers déposés ainsi que leurs copies d'archivage ne sont pas protégés par des mesures techniques de protection (DRM).