



**HAL**  
open science

# Revue académiques : nouvelles opportunités pour la visibilité des articles. Le cas de la diffusion des métadonnées de la Revue scientifique et technique de l'OIE

Romuald Verrier

## ► To cite this version:

Romuald Verrier. Revue académiques : nouvelles opportunités pour la visibilité des articles. Le cas de la diffusion des métadonnées de la Revue scientifique et technique de l'OIE. domain\_shs.info.docu. 2015. mem\_01309538

**HAL Id: mem\_01309538**

[https://memic.ccsd.cnrs.fr/mem\\_01309538v1](https://memic.ccsd.cnrs.fr/mem_01309538v1)

Submitted on 29 Apr 2016

**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.



Distributed under a Creative Commons Attribution 4.0 International License

CONSERVATOIRE NATIONAL DES ARTS ET METIERS

Ecole Management et Société-Département CITS

INTD

MEMOIRE pour obtenir le Titre enregistré au RNCP

"Chef de projet en ingénierie documentaire"

Niveau I

Présenté et soutenu par

*Romuald Verrier*

le 29 octobre 2015

Revue académiques : nouvelles opportunités  
pour la visibilité des articles.

Le cas de la diffusion des métadonnées  
de la *Revue scientifique et technique* de l'OIE

Jury :

Madame Ghislaine Chartron, directrice du département Culture, Information,  
Technique et Société de l'INTD-CNAM

Madame Aline Rousier, documentaliste de l'OIE

**Promotion 45**

## Remerciements

C'est un grand sentiment de reconnaissance qui m'anime aujourd'hui à l'égard de Madame Aline Rousier, qui m'a accueilli avec bienveillance, m'a donné l'opportunité de réaliser cette étude, et m'a continuellement encouragé à avancer.

Je tiens à remercier tout particulièrement le Docteur Daniel Chaisemartin, qui m'a soutenu chaleureusement et qui m'a accordé sa confiance. Il m'a ouvert les portes de son service dans un esprit de transparence, de disponibilité et de clarté sur les objectifs à atteindre.

Je remercie aussi Madame Ghislaine Chartron, pour ses conseils et encouragements.

Monsieur Bertrand Flahaut m'a beaucoup aidé par son écoute, ses conseils et ses avis techniques avisés. Je le remercie sincèrement.

Je n'oublie pas l'équipe du service des publications de l'OIE, qui m'a fait partager sa bonne humeur et son professionnalisme. Par sa disponibilité et ses conseils, elle a largement contribué à la richesse des connaissances acquises au cours du projet.

Enfin, j'adresse mes remerciements à toutes les personnes – à l'OIE mais aussi chez la société GB Concept, à l'ABES, chez CrossRef et PubMed – qui ont accepté de m'accorder un temps souvent précieux pour répondre à mes questions.

## Notice

VERRIER Romuald. Revues académiques numériques : nouvelles opportunités pour la visibilité des articles. Le cas de la diffusion des métadonnées de la Revue scientifique et technique de l'OIE. 2015. X pages. Mémoire pour obtenir le titre professionnel « Chef de projet en ingénierie documentaire », CNAM-INTD, 2015.

**Résumé :** Ce mémoire examine comment l'exploitation d'outils de gouvernance des métadonnées permet d'accroître la présence en ligne et la visibilité d'une revue académique numérique. Après un bref aperçu des évolutions récentes dans le monde des revues académiques, l'auteur présente la façon dont la *Revue scientifique et technique de l'OIE* peut bénéficier de l'interopérabilité des métadonnées en s'appuyant sur son portail documentaire. L'auteur examine l'impact des moteurs de recherche, des bases d'indexation, des bases de connaissances, des outils de citation et des réseaux sociaux professionnels, et présente la mise en œuvre de solutions : SEO, DOI, flux XML, OAI-PMH, KBART et politique de libre accès. Ce mémoire pourra intéresser les éditeurs, bibliothécaires, intermédiaires commerciaux et tout professionnel confronté aux métadonnées de revue académique.

**Descripteurs :** Accès libre, Alexandrie, Base de connaissances, Bibliométrie, DOI, Donnée bibliographique, Dublin Core, Information scientifique et technique, Ingénierie documentaire, Interopérabilité, KBART, Métadonnée, OAI-PMH, OIE, Périodique, Portail documentaire, PubMed, Recommandation, Référence bibliographique, Référencement, Revue électronique, SEO, Structuration de données, Visibilité, XML, XSLT.

**English summary :** This study examines how to exploit the full potential of metadata management tools to expand the online presence and visibility of an online journal. After a short overview of the current trends in the world of scholarly journals, the author shows how the Scientific and technical review of the OIE can benefit from metadata interoperability based on its documentation portal. The author examines the impact of search engines, A&I services, knowledge bases, citation management tools and popular academic social network sites and presents the implementation of a few solutions, ranging from SEO, DOI, XML feeds, OAI-PMH, KBART to open access policies. This paper will be of value to publishers, librarians, vendors and to any professional dealing with journal metadata.

**Keywords :** Abstract and indexing services, Bibliometrics, Citation Manager, Citations, DOI, Dublin Core, E-journals, Google Scholar, Journals, KBART, Knowledge bases, Metadata, OAI-PMH, OIE, Open access, Peer-review, PubMed, Recommender Systems, SEO, STM, XML, XSLT.

# Table des matières

Remerciements.....	1
Notice.....	2
Table des matières .....	4
Liste des tableaux.....	7
Liste des figures .....	8
<b>Introduction .....</b>	<b>10</b>
<b>1 Les revues académiques : définition et données clés .....</b>	<b>14</b>
1.1 Panorama mondial des revues académiques en STM.....	14
1.2 Évolutions récentes.....	16
1.2.1 La numérisation .....	16
1.2.2 Le libre accès .....	17
1.2.3 Vers de nouvelles mesures statistiques .....	19
<b>2 Améliorer la visibilité de la <i>Revue scientifique et technique</i> de l'OIE : quels besoins ? .....</b>	<b>21</b>
2.1 Présentation de la démarche .....	21
2.1.1 Le projet de stage .....	21
2.1.2 Méthodologie utilisée .....	23
2.2 Le contexte du projet : éléments clés.....	25
2.2.1 L'OIE.....	25
2.2.2 La <i>Revue scientifique et technique</i> .....	26
2.2.3 Le Service de l'administration, de la logistique et des publications de l'OIE 28	
2.2.3.1 Les systèmes d'information .....	29
2.2.3.2 L'édition .....	30
2.2.3.3 La documentation .....	31
2.2.3.3.1 La base documentaire .....	31

2.2.3.3.2	Le logiciel Alexandria .....	33
2.3	Les objectifs de visibilité : priorités .....	34
2.3.1	Les moteurs de recherches .....	35
2.3.2	Les bases d'indexation .....	37
2.3.2.1	MEDLINE/PubMed – NCBI .....	38
2.3.2.2	CAB Abstracts – CABI .....	40
2.3.2.3	Web of Knowledge – Thomson Reuters.....	41
2.3.2.4	SCOPUS – Elsevier.....	43
2.3.2.5	AGRICOLA – US National Agricultural Library .....	44
2.3.2.6	DOAJ.....	45
2.3.2.7	AGRIS – FAO .....	46
2.3.2.8	HINARI – OMS .....	47
2.3.2.9	The Capsule Report.....	48
2.3.2.10	Bases abandonnées .....	48
2.3.2.10.1	Focus On: Veterinary Science and Medicine.....	48
2.3.2.10.2	Index Medicus.....	49
2.3.2.10.3	Littérature vétérinaire francophone.....	49
2.3.2.10.4	PASCAL – INIST .....	49
2.3.3	Les catalogues de bibliothèques et les bases de connaissance .....	50
2.3.4	Les outils de citations et les réseaux sociaux académiques, scientifiques et professionnels.....	54
2.3.4.1	Les outils de citations .....	54
2.3.4.2	Les réseaux sociaux académiques, scientifiques et professionnels ....	56
2.4	Conclusion intermédiaire .....	60
<b>3</b>	<b>Améliorer la visibilité de la <i>Revue scientifique et technique</i> de l'OIE : quelles solutions ?.....</b>	<b>61</b>
3.1	Optimisation du référencement naturel .....	61
3.1.1	Architecture et navigation au sein du site internet de l'OIE .....	62
3.1.2	Instructions aux robots d'indexation .....	68
3.1.3	Enrichissement sémantique des pages du portail documentaire .....	68

3.2	Identifiants pérennes .....	71
3.2.1	DOI .....	72
3.2.1.1	Comparaison des agences CrossRef et DataCite.....	75
3.2.1.2	Travailler avec CrossRef.....	80
3.2.1.2.1	Calendrier prévisionnel .....	80
3.2.1.2.2	Principaux services impactés .....	81
3.2.1.2.3	Dépôts XML .....	82
3.2.2	PMID (PubMed ID) .....	83
3.3	Exports XML.....	84
3.3.1	Remarques sur la DTD PubMed .....	89
3.3.2	Remarques sur le schéma CrossRef .....	91
3.4	Entrepôt OAI-PMH .....	91
3.4.1	Le protocole OAI-PMH .....	91
3.4.2	Le format Dublin Core.....	92
3.4.3	Installation d'un entrepôt OAI-PMH à l'OIE.....	94
3.5	Fichiers KBART .....	99
3.6	Positionnement sur le libre accès .....	101
	<b>Conclusion .....</b>	<b>110</b>
	<b>Bibliographie .....</b>	<b>113</b>
	Annexe 1 Citations croisées avec CrossRef.....	123



## Liste des tableaux

TABLEAU 1 : FACTEURS D'IMPACT DE LA <i>REVUE SCIENTIFIQUE ET TECHNIQUE</i> DE L'OIE POUR LA PERIODE 2008-2014 (SOURCE : JCR).....	27
TABLEAU 2 : LES QUATRE GRANDES NAVIGATIONS POSSIBLES POUR LA RECHERCHE D'INFORMATION EN LIGNE, EN SCIENCES VETERINAIRES ET EN SCIENCES BIOLOGIQUES, ET LES LEVIERS D'ACTION POUR Y ACCROITRE LA VISIBILITE DE LA <i>REVUE SCIENTIFIQUE ET TECHNIQUE</i> DE L'OIE.....	35
TABLEAU 3 : COMPARAISON DES AGENCES DOI DATA CITE ET CROSSREF .....	76
TABLEAU 4 : EQUIVALENCE EN DUBLIN CORE DES ETIQUETTES DE CHAMPS ALEXANDRIE.....	85

## Liste des figures

FIGURE 1 : ORGANIGRAMME SIMPLIFIE DU SIEGE DE L'OIE (SOURCE : OIE) .....	29
FIGURE 2 : REPARTITION DES PROFILS PAR DISCIPLINES AU SEIN DES QUATRE PRINCIPAUX RESEAUX SOCIAUX ACADEMIQUES .....	56
FIGURE 3 : NAVIGATION POUR ACCEDER AU TEXTE INTEGRAL D'UN ARTICLE DE LA <i>REVUE SCIENTIFIQUE ET TECHNIQUE</i> (ETAT AU 31 AOUT 2015) .....	65
FIGURE 4 : NAVIGATION POUR ACCEDER AU TEXTE INTEGRAL D'UN ARTICLE DE LA <i>REVUE SCIENTIFIQUE ET TECHNIQUE</i> APRES MISE EN LIGNE DU NOUVEAU PORTAIL DOCUMENTAIRE DE L'OIE .....	66
FIGURE 5 : METADONNEES BIBLIOGRAPHIQUES IDENTIFIEES PAR L'OUTIL DE CITATION ZOTERO A PARTIR DES BALISES <i>META</i> DE PAGES HTML. COMPARAISON DES DONNEES IDENTIFIEES SUR LA NOTICE D'UN ARTICLE DANS LA PARTIE MARCHANDE DU SITE (A GAUCHE, PAGE SANS BALISE META) ET SUR LA NOTICE EQUIVALENTE DANS LE PORTAIL .....	70
DOCUMENTAIRE (A DROITE, APRES AJOUT DES SCHEMAS DUBLIN CORE, CITATION_ ET PRISM DANS LES BALISES <i>META</i> ) .....	70
FIGURE 6 : BALISES <i>META</i> D'UNE NOTICE SUR LE PORTAIL DOCUMENTAIRE DE L'OIE (AFIN DE GARANTIR UNE BONNE LISIBILITE, IL S'AGIT D'UNE COPIE ECRAN SANS RETOUR A LA LIGNE AUTOMATIQUE, CE QUI PERMET DE VOIR L'ENSEMBLE DES BALISES MALGRE LE TEXTE IMPORTANT CONTENU PAR EXEMPLE DANS LES RESUMES) .....	71
FIGURE 7 : EXTRAIT DU FICHIER XSLT POUR LES FLUX XML VERS PUBMED .....	90
FIGURE 8 : EXEMPLE DE NOTICE RECUPERABLE EN OAI-PMH .....	97
FIGURE 9 : EXEMPLE D'UNE POLITIQUE DE DEPOT D'ARTICLES SUR LE SITE HELOÏSE .....	107
FIGURES 10 ET 11 : EXEMPLES D'UNE POLITIQUE DE DEPOT D'ARTICLES SUR LE SITE SHERPA ROMEO .....	108
FIGURE 12 : CODES COULEUR DES POLITIQUES DE DEPOT D'ARTICLES SUR LE SITE SHERPA ROMEO .....	109
FIGURE 13 : EXEMPLE DE FICHE EDITEUR SUR LE SITE DOAJ .....	109



# **Introduction**

Dans l'univers du livre, le codex fait l'objet d'un certain nombre de conventions, avec des formats et des gabarits privilégiés selon les thèmes traités, le public ciblé, etc. La forme des revues académiques imprimées était également très codifiée par les usages, avec une architecture générale prévisible, fruit d'une certaine convention de lecture entre les revues et leur public. Passées quelques velléités expérimentales au milieu des années quatre-vingt-dix, les usages ont à nouveau convergé pour la numérisation des revues académiques, entraînant l'émergence de bonnes pratiques, normes et standards nouveaux.

Cette convergence est certainement une conséquence de la forte concentration du secteur [4, WARE, p. 45]. D'autre part, depuis des années, les acteurs de l'interprofession (éditeurs, associations de bibliothécaires, producteurs de bases d'indexation, agences d'abonnements, éditeurs d'outils de découverte, etc.) se rencontrent et échangent dans de nombreux groupes de travail internationaux. Il s'agit notamment d'offrir de nouveaux services, fondés sur les opportunités du numérique.

Un des sujets les plus discutés concerne les données descriptives des revues et de leurs articles (on parle alors de métadonnées : titre, auteur(s), langue, date de publication, résumé, format, droits associés, etc.).

Pour les éditeurs, une opportunité nouvelle concerne l'optimisation de la visibilité de leurs ressources numériques. Les nombreux travaux sur l'interopérabilité des métadonnées ont abouti à des protocoles, normes ou recommandations, qui développent considérablement la qualité du référencement et la « trouvabilité » des articles. Il apparaît aujourd'hui que des actions modestes peuvent avoir un impact fort sur la visibilité d'une revue.

Sur ces bases, l'OIE<sup>1</sup>, Organisation mondiale de la santé animale, éditrice de la *Revue scientifique et technique*, s'interrogeait sur la meilleure façon d'améliorer la visibilité des articles de la revue.

En effet, à l'occasion de la migration de la base documentaire de l'OIE vers une nouvelle version du logiciel Alexandrie (société GB Concept), il était possible de mettre en œuvre une nouvelle gouvernance des métadonnées. Quelles solutions complémentaires étaient disponibles pour améliorer l'interopérabilité des données ? Quels formats choisir ? Où et comment les exposer ? Avec quels objectifs ?

Dans le cadre d'un stage effectué entre juin et septembre 2015, le Chef du Service de l'administration, de la logistique et des publications m'a confié une recherche sur l'exposition des métadonnées de la *Revue scientifique et technique*, ce qui incluait la mise en œuvre de solutions et la rédaction d'un certain nombre de notes, soit à vocation descriptive sur les actions conduites, soit à titre de recommandations.

Après un panorama des évolutions récentes du secteur des revues académiques, ce mémoire examine comment l'exploitation d'outils de gouvernance des métadonnées peut accroître la présence en ligne et la visibilité de la *Revue scientifique et technique* de l'OIE.

La deuxième partie est consacrée à l'analyse des stratégies de référencement pertinentes dans le contexte de l'OIE. Pour cela, j'ai étudié les navigations possibles d'une recherche documentaire en sciences vétérinaires (moteurs de recherche, bases d'indexation, bases de connaissance ou outils de découverte, outils de citation ou réseaux sociaux professionnels) et j'ai recommandé pour chacune les stratégies les mieux adaptées.

---

<sup>1</sup> L' OIE est une organisation intergouvernementale de 180 pays membres – [www.oie.int](http://www.oie.int)

La troisième et dernière partie de ce mémoire décrit les leviers d'action recommandés et leur mise en œuvre à l'OIE : référencement naturel, identifiants pérennes, exports XML, OIA-PMH, KBART et politique de libre accès.

# 1 Les revues académiques : définition et données clés

---

## 1.1 Panorama mondial des revues académiques en STM

Sous le terme générique de revue académique, on désigne le spectre assez large des publications périodiques savantes dotées d'un comité de lecture ou d'un comité scientifique, qu'elles s'adressent à un public universitaire ou davantage à des praticiens, aussi bien en sciences, techniques et médecine (STM) qu'en sciences humaines et sociales (SHS). Notre analyse porte essentiellement sur les revues STM puisque le marché et les modèles économiques en SHS sont fortement dépendants des contextes locaux ou linguistiques.

L'association STM (< <http://www.stm-assoc.org/> >) estime qu'il existe 34 550 revues académiques actives dans le monde en 2014 (dont 28 100 en langue anglaise). Ces revues génèrent 2,5 millions d'articles par an (taux de croissance annuel compris entre 3 % et 3,5 %) rédigées par 7 à 8 millions de chercheurs-auteurs (taux de croissance annuel également à 3 %) [4, WARE, p. 6].

Une majorité de sociétés savantes délègue par appels d'offres l'édition et la promotion de leurs revues ; un petit nombre d'éditeurs se partage le marché. Ce secteur est en effet très fortement concentré puisque les 10 premiers éditeurs mondiaux possèdent 45,2 % des revues. Chacun des quatre premiers éditeurs (Springer-Nature Publishing Group, Elsevier, Wiley, Taylor & Francis) publient plus de 2000 titres [4, WARE, p. 45].

Dans leur ensemble, les revues académiques jouent un rôle clé dans la recherche. Elles interviennent essentiellement à la fin d'un cycle qui peut être synthétisé en quatre phases :

- 1. découvertes, hypothèses
- 2. financement, approbation
- 3. recherche (protocole, résultats, validation)



- 4. dissémination des résultats

Le rôle des revues pour la recherche peut encore être précisé. On attribue à Zuckerman et Merton d'avoir énoncé les quatre fonctions des revues académiques :

- enregistrer (l'antériorité d'une découverte ou d'une idée, la propriété d'un écrit) ;
- disséminer (et contextualiser auprès d'un public ciblé, au travers d'une stratégie de marque) ;
- certifier (par la caution scientifique d'une revue par les pairs) ;
- archiver (préservé une version définitive de l'article qui pourra être citée, référencée).

Par l'évaluation de la qualité des travaux académiques, les revues jouent un rôle de filtre face à l'abondance des recherches. Ce rôle des revues dans la validation des résultats de la recherche génère des stratégies de marque mais avec de grandes différences selon les domaines, comme le note Olivier Bomsel (MINES ParisTech – CERNA) :

« Plus une marque attire d'auteurs de qualité, plus est forte l'utilité qu'elle délivre à ses lecteurs, et réciproquement. Si la qualité obéit à des critères objectifs – de validité ou de portée des résultats, par exemple –, de telles plate-formes ont vocation à devenir des monopoles sur chaque marché où elles opèrent. C'est le cas des revues de sciences techniques et médicales (STM), dont, en raison de la spécialisation des chercheurs, les marchés sont ultra-segmentés. Plus interdisciplinaires, moins objectivables, les sciences humaines et sociales (SHS) s'éditent également à travers des ouvrages : le pouvoir de marché des marques éditoriales y est donc moins flagrant. »  
[14, BOMSEL, p. 99]

## 1.2 Évolutions récentes

Si l'histoire des revues académiques est riche de plusieurs siècles (le *Journal des sçavans*, dont le premier numéro date de 1665, est considéré comme le plus ancien périodique littéraire et scientifique d'Europe), les changements intervenus depuis les années quatre-vingt-dix ont profondément modifiés le secteur, sur un temps au final relativement court. Ainsi, sous le double effet de la numérisation et de l'émergence du libre accès, les revues académiques ont dû se transformer et s'adapter rapidement, dans un paysage en évolution permanente.

### 1.2.1 La numérisation

Au cours des siècles, le format des revues s'était relativement standardisé, par exemple avec des informations d'identification prévisibles et un certain nombre de conventions de présentation.

Avec les revues numériques, les éditeurs ont parfois été tenté d'innover de façon indépendante (sur les formats, la présentation des titres, les formules d'abonnements...). Cette période fut brève et, depuis plusieurs années, il existe à l'inverse une convergence forte des modèles, et une volonté consensuelle de travailler sur l'émergence de standards communs. Il s'agit par exemple d'améliorer la qualité des métadonnées, la « trouvabilité » des articles dans les bases de connaissance et les moteurs de recherche, et d'ouvrir les revues à de nouveaux usages.

Le guide de bonnes pratiques *PIE-J The Presentation & Identification of E-Journals* [16, NISO] illustre parfaitement ce phénomène. Au travers des nombreux exemples de sites internet de revues académiques, dont la structuration fait l'objet d'une analyse avancée, le guide démontre la convergence des pratiques considérées comme meilleures, car plus efficaces en terme de visibilité. Une revue ne peut ignorer cet écosystème global, dont certains usages communs sont autant de prescriptions impératives pour conserver un certain niveau de visibilité et de crédibilité.

L'environnement numérique a modifié bien des modèles économiques, en remplaçant par un coût fixe au premier exemplaire le coût variable des exemplaires imprimés. Certains éditeurs ont pris prétexte des lourds investissements liés au numérique pour augmenter exagérément les tarifs d'abonnement. Cet opportunisme fut souvent très mal accueilli par un certain nombre d'acteurs, dont beaucoup s'engagèrent fermement dans le mouvement alors émergent de l'open access.

### 1.2.2 Le libre accès

Pour une revue académique, être en libre accès (*open access*) signifie la mise à disposition gratuites des articles numérisés, avec une licence ouverte et sans obstacle technique à l'accès. Selon l'association STM, une politique de libre accès doit définir le **quoi**, le **quand** et le **comment** de ce qui est ouvert [4, WARE, p. 88].

#### Quoi

Il s'agit de préciser ce qui est ouvert, c'est-à-dire quelle version du texte est disponible :

- version soumise ;
- version acceptée ;
- version publiée.

L'organisation NISO propose encore davantage d'états : *Author's Original, Submitted Manuscript Under Review, Accepted Manuscript, Proof, Version of Record, Corrected Version of Record, Enhanced Version of Record...*).

#### Quand

La date de disponibilité en libre accès des articles est soit :

- avant publication ;
- immédiatement après publication ;
- passé un délai (période dite d'embargo) après publication.

## Comment

Enfin, les trois principaux modèles de publication/diffusion en Open Access (OA) sont :

- la voie dorée (*Gold OA*) : la version publiée est librement accessible, immédiatement après publication ;
- le libre accès avec embargo : la version publiée devient librement accessible après une période d'embargo ;
- la voie verte (*Green OA*) : la version acceptée est auto-archivée par l'auteur, immédiatement ou après publication. Parmi les réservoirs d'auto-archivage les plus connus, citons :
  - HAL < <https://hal.archives-ouvertes.fr/> > (réservoir français maintenu par le CNRS) ;
  - ArXiv < <http://arxiv.org/> > (maintenu par la bibliothèque de l'université Cornell) ;
  - PubMed Central < <http://www.ncbi.nlm.nih.gov/pmc/> > (projet issu du site PubMed – dont il bénéficie de la notoriété –, spécialisé en médecine et en sciences biologiques) ;
  - SSRN < <http://www.ssrn.com/en/> > (Social Science Research Network, spécialisé en sciences sociales).

Une cartographie intéressante des modèles de libre accès est intitulée *How Open Is It ?* (< [http://www.sparc.arl.org/sites/default/files/hoii\\_guide\\_rev4\\_web.pdf](http://www.sparc.arl.org/sites/default/files/hoii_guide_rev4_web.pdf) >, consulté le 15 octobre 2015). Proposée par l'association SPARC < [www.arl.org/sparc](http://www.arl.org/sparc) > et l'éditeur PLOS < [www.plos.org](http://www.plos.org) >, son originalité repose sur l'analyse de 6 critères décomposés en 4 ou 5 degrés d'ouverture : droits du lecteur (ou de l'abonné), droits de réutilisation (notamment au travers des licences Creative Commons), copyrights, droits de l'auteur (par exemple pour l'auto-archivage et publication des articles sur d'autres sites), archivage automatique et lisibilité par les

machines (accès aux métadonnées, API, fouille de données et *text-mining*...). L'ensemble des combinaisons possibles illustre bien la diversité des modèles existants.

Le modèle de la voie dorée transforme radicalement l'économie des revues. Puisque le contenu est librement accessible pour les lecteurs alors l'activité éditoriale devient en quelque sorte un service aux auteurs. Afin de faire rémunérer leurs services, les éditeurs appliquent donc des APC (*article processing charges*), sur le principe de l'« auteur-payeur ». Les discussions sur les vertus et inconvénients de ce modèle sont souvent passionnées car lourdes d'enjeux, comme en témoigne une abondante littérature. Ce débat dépasse cependant largement le cadre du présent mémoire, bien qu'une politique de libre accès soit un levier possible de visibilité.

En décembre 2014, l'annuaire DOAJ (Directory of Open Access Journals, < <https://doaj.org> >) référençait 10 091 revues en libre accès. Chiffres qu'il est intéressant de nuancer avec les indications de la base d'indexation Scopus, qui n'identifie que 13 % de périodiques open access parmi les 22 000 revues qu'elle référence. Une autre très grande base d'indexation, Web of Science (de Thomson Reuters) ne qualifie comme open access que 9 % des revues suivies [4, WARE, p. 31].

### **1.2.3 Vers de nouvelles mesures statistiques**

Être publié permet à l'auteur de disséminer les résultats de sa recherche et, entre autres motivations, de favoriser son avancement de carrière et le financement de ses recherches futures. Mais, au-delà de la publication, les bailleurs de fonds de la recherche souhaitent des preuves de la performance de leur investissement. Ils apprécient souvent que les travaux soient validés également par d'autres sources, telles que les nouveaux outils d'évaluation et de bibliométrie (usages, citations, partages sur les réseaux sociaux).

L'harmonisation des formats de métadonnées, la convergence des protocoles d'interopérabilité ont en effet permis de quantifier la performance au niveau des articles, et non plus seulement au niveau des revues.

Ainsi, le protocole SUSHI et le format COUNTER 4, établis par l'agence de normalisation NISO, mesurent les usages de consultation en bibliothèque.

De même, des outils bibliométriques récents tels que Altmetrics < <http://altmetrics.org> > ou PlumX < <http://plumanalytics.com> > analysent l'impact des articles au-delà leurs citations dans d'autres articles : partage sur les réseaux sociaux, nombre de téléchargements, recommandations...

La valeur de contextualisation des éditeurs pourrait tendre alors à s'estomper, au risque que l'article devienne une unité isolée, autonome. Ce ne serait alors plus la marque éditoriale qui déterminerait la visibilité d'un texte mais la pertinence de son référencement par les moteurs de recherche, sa popularité au sein d'une communauté active ou encore la froide recommandation algorithmique...

## 2 Améliorer la visibilité de la *Revue scientifique et technique* de l'OIE : quels besoins ?

---

### 2.1 Présentation de la démarche

#### 2.1.1 Le projet de stage

Le projet visait à étudier les opportunités induites par un nouvel outil documentaire pour la gestion et la diffusion des métadonnées des publications de l'OIE. Plus particulièrement, le Docteur Daniel Chaisemartin, Chef du Service de l'administration, de la logistique et des publications de l'OIE, et Aline Rousier, documentaliste, souhaitaient explorer les avantages d'une centralisation des données dans un outil unique. La question posée était : « comment exposer et/ou diffuser les métadonnées administrées dans le nouvel outil pour favoriser une meilleure visibilité des publications de l'OIE – en priorité la *Revue scientifique et technique* – et la base de données documentaire de l'OIE ? » Un objectif tacite mais de même niveau était que la consultation des articles devait avoir lieu sur le site de l'OIE afin, d'une part, d'y agréger et d'y mesurer l'audience, de permettre une contextualisation éventuelle des documents, et d'autre part, d'informer au mieux le lecteur sur les activités de l'OIE.

J'ai traité cette question dans le cadre du stage effectué à l'OIE, sous la responsabilité d'Aline Rousier, de juin à août 2015. Une première reformulation de l'objectif final avait permis d'identifier un certain nombre d'exigences déjà formulées. Ainsi, l'étude devait porter en priorité sur la visibilité de la *Revue scientifique et technique*, traiter des bases d'indexation PubMed et CABI, évaluer l'intérêt d'identifiants pérennes de type DOI. Et, au-delà de recommandations à fonder sur les normes, bonnes pratiques

ou usages des revues académiques, il s'agissait également de mettre en œuvre un entrepôt OAI-PMH et de rendre les métadonnées interoperables.

Certains de ces éléments étaient déjà clairement formulés dans une présentation interne, rédigée à l'automne 2014, portant sur la visibilité de la base de données de l'OIE et de la librairie en ligne (partie marchande du site internet de l'OIE). Cette présentation – que je ne reproduirai pas ici par souci de confidentialité – avait pour finalité d'ouvrir une discussion au sein du Service de l'administration, de la logistique et des publications de l'OIE sur plusieurs pistes qu'ils convenaient d'explorer en matière de visibilité. Le cadre du stage avait donc été particulièrement bien réfléchi en amont de mon arrivée. Plus précisément, comme le stage recouvrait d'autres missions connexes, ce projet-ci nécessitait d'étudier les opportunités d'exposition des métadonnées afin d'aboutir à :

- des recommandations sur les décisions à prendre ;
- de brefs documents descriptifs, de type « manuels », pour les solutions envisagées ;
- des réalisations effectives (et documentées) pour les solutions validées.

Ce programme fut respecté et le présent mémoire s'appuie en grande partie sur les travaux effectués dans le cadre du stage.

Cette deuxième partie décrit et explicite les besoins identifiés, et nomme les solutions recommandées. La troisième et dernière partie du mémoire détaillera plus longuement les solutions retenues et mises en œuvre, en cours de déploiement ou restant à arbitrer, et leurs impacts sur l'organisation de l'OIE.



### **2.1.2 Méthodologie utilisée**

Dans un premier temps, afin d'éviter tout risque d'omission dans la définition du périmètre du projet, j'ai appliqué la méthode de questionnement systématique et exhaustif dite « QQQQCCP » (pour « Qui ? Quoi ? Où ? Quand ? Comment ? Combien ? Pourquoi ? »).

Pour dégager les objectifs et concepts du projet, je me suis appuyé sur une analyse systématique des documents internes (présentations, comptes rendus de réunions, correspondances du service Documentation) ou externes (documentation du logiciel Alexandria de la société GB Concept) et avons conduit des entretiens internes répartis sur les deux premières semaines (chef de service, documentaliste, chef de la Cellule des publications, responsable systèmes d'information, gestionnaire de ventes et marketing).

Pour le premier entretien avec la documentaliste de l'OIE, j'ai repris le questionnaire utilisé par l'ABES (Agence bibliographique de l'enseignement supérieur) pour son « Enquête sur les métadonnées dans l'édition numérique (projet BACON) » réalisée entre le 3 et 31 mars 2015 au travers d'un questionnaire en ligne (< [https://fr.surveymonkey.com/s/ABES\\_BACON](https://fr.surveymonkey.com/s/ABES_BACON) >, consulté le 15 septembre 2015). Cette enquête visait à recueillir, auprès d'éditeurs de revues académiques, des informations sur l'exposition des métadonnées relatives à leurs publications numériques. Le questionnaire était donc idéalement adapté à l'objectif général défini par l'OIE. Ainsi, j'ai pu évaluer facilement l'existant, identifier les parties prenantes et leur rôle, et appréhender les principales contraintes et exigences.

Les entretiens qui ont suivi ont permis d'approfondir les pratiques et les besoins propres à chaque partie.

Après une première recherche documentaire, j'ai mis en œuvre une veille thématique bilingue (français et anglais) sur la visibilité des revues académiques. Les principaux

sujets suivis étaient les identifiants pérennes, les formats de métadonnées, les politiques de libre accès, les bases de connaissance et les outils de découverte, les réseaux sociaux académiques et les outils de citation.

Parallèlement, j'ai effectué une analyse comparée (*benchmark*) avec plusieurs revues, parmi lesquelles :

- *Onderstepoort Journal of Veterinary Research* < <http://www.ojvr.org> >

- *Journal of Virology* (revue de l'American Society for Microbiology),  
< <http://jvi.asm.org> >

- *Revue Internationale du Froid* (revue de l'Institut international du Froid, éditée par Elsevier), < [www.elsevier.com/locate/ijrefrig/](http://www.elsevier.com/locate/ijrefrig/) >

- *Parasite* (revue de la Société Française de Parasitologie, éditée par EDP Sciences),  
< <http://www.parasite-journal.org> >

Le choix de ces revues était fondé sur les conseils de la documentaliste de l'OIE et sur des critères de similarité des problématiques de visibilité par rapport à la *Revue scientifique et technique*. Cette analyse m'a permis de recueillir des informations précieuses sur les usages en matière de stratégies de référencement, de multilinguisme, de structuration des pages Web des articles, de politique de libre accès.

L'ensemble des éléments recueillis a, entre autres choses, permis d'établir une liste des stratégies de recherche d'information mises en œuvre par les lecteurs-cibles de la *Revue scientifique et technique* de l'OIE. C'est en effet en partant des usages que j'ai identifié les leviers d'action permettant, dans le contexte propre à l'OIE, d'atteindre les objectifs de visibilité attendus.

Sur ces bases, j'ai transmis une série de notes internes au Service de l'administration, de la logistique et des publications de l'OIE. Leur contenu présentait la démarche poursuivie, les recommandations et le détail des réalisations effectuées. Il convenait en effet d'explicitier le choix des outils et de documenter leur mise en œuvre (effective ou à venir) par l'OIE.

Cette deuxième partie du mémoire reprend et approfondit le choix des outils. Leur mise en œuvre et les contraintes de déploiement sont présentées dans la troisième partie, qui complète cette deuxième partie.

## **2.2 Le contexte du projet : éléments clés**

### **2.2.1 L'OIE**

L'Organisation mondiale de la santé animale (OIE) est chargée d'améliorer la santé et le bien-être animal dans le monde. Cette organisation inter-gouvernementale compte 180 pays membres, et a son siège à Paris.

L'Office international des épizooties (OIE) fut créé en 1924, après l'apparition de la peste bovine en Europe. En mai 2003, l'Office est devenu l'Organisation mondiale de la santé animale, mais l'acronyme historique OIE a été conservé.

L'OIE gère le système mondial de veille et d'alerte zoonositaire et joue un rôle clé dans les domaines de la recherche et de l'information scientifique vétérinaire. Elle élabore les normes sanitaires qui encadrent le commerce mondial des animaux et de leurs produits et assurent la sécurité sanitaire des aliments d'origine animale dès leur production. L'OIE est également l'unique organisation mondiale pour le bien-être animal.

### **2.2.2 La *Revue scientifique et technique***

La *Revue scientifique et technique* de l'OIE est une publication périodique trilingue (français, anglais et espagnol : les trois langues officielles de l'organisation), disponible en version imprimée (ISSN 0253-1933) et en version numérique (ISSN 1608-0637).

Comme précisé sur le site internet de l'OIE, la *Revue scientifique et technique* répond à certains objectifs prioritaires de l'organisation. Évaluée par des pairs au travers d'un Comité consultatif de rédaction et d'un Comité scientifique et technique, elle diffuse des articles reflétant les dernières avancées scientifiques et techniques en matière de santé animale, de santé publique vétérinaire, de sécurité sanitaire des aliments et de bien-être animal dans le monde.

Les articles soumis en français et en espagnol sont traduits systématiquement en anglais et les deux versions de l'article sont publiées dans la revue. Les résumés des articles sont, par contre, toujours publiés dans les trois langues.

Publiée depuis 1982 au rythme de trois numéros par an (dont deux numéros thématiques – avec des articles demandés à des auteurs – et un troisième numéro plurithématique – avec des articles soumis spontanément par leurs auteurs –), la version imprimée est disponible par abonnement annuel ou par achat au numéro. Chaque numéro comprend entre 250 et 350 pages environ, en quadrichromie et au format 21 x 29,7 cm. Le tirage de l'édition imprimée varie autour de 1500 exemplaires par numéro, dont 400 sont adressés gratuitement aux partenaires et Délégués de l'OIE.

Les articles sont disponibles en texte intégral (format PDF de la version publiée) sur le site internet de l'OIE, en accès libre mais avec un décalage de quelques mois à partir de la parution de la version imprimée.

L'équipe éditoriale est propre à l'OIE, intégrée au Service de l'administration, de la logistique et des publications de l'OIE, et placée sous la responsabilité d'Annie Souyri, chef de la Cellule des publications, et de Daniel Chaisemartin, chef du Service.

En raison du programme éditorial, avec la définition d'un sommaire après sélection d'un coordonnateur du numéro de la revue puis demande d'articles à des auteurs proposés par le coordonnateur de la revue, chaque numéro thématique est initié très en amont de sa publication effective (jusqu'à trois ans).

Pour cette raison également, le classement de la revue par les outils bibliométriques n'est pas regardé en interne comme un indicateur pertinent. Malgré cela, et malgré les critiques régulièrement adressées à ces outils par la profession, il est important de noter les bonnes statistiques de citation relevées par le facteur d'impact (*Impact factor*). Créé en 1955 par Eugène Garfield de la société Thomson Reuters qui publie chaque année le Journal Citations Reports (JCR) avec les facteurs d'impact des revues indexées dans la base de données Web of Science (WoS), l'Impact factor compare le nombre de citations reçues en l'an X avec l'ensemble des articles publiés dans la revue au cours des deux années précédentes, rapporté au nombre d'articles publiés par la revue pendant la même période.

**Tableau 1 : Facteurs d'impact de la *Revue scientifique et technique* de l'OIE pour la période 2008-2014 (source : JCR)**

Année	Impact Factor (IF)	Nombre d'articles
2013/2014	0.910	85
2012	0.69	46
2011	1.099	77
2010	1.609	49
2009	1.238	75
2008	0.918	67

D'autres outils bibliométriques célèbres quantifient également le « poids » de la revue et/ou des articles qui y sont publiés. Ainsi, Elsevier propose avec le SCImago Journal & Country Rank une vision complémentaire intéressante (< <http://www.scimagojr.com/journalsearch.php?q=0253-1933&tip=iss> >, consulté le 15 septembre 2015).

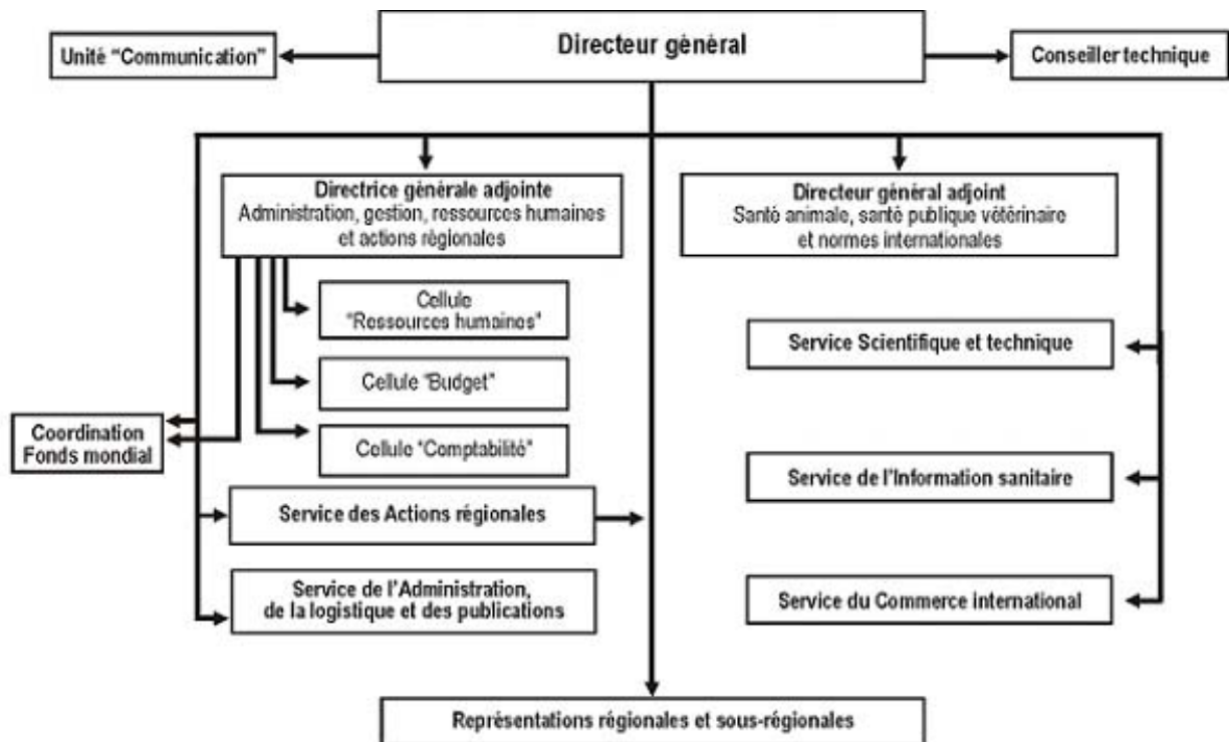
L'OIE n'utilise pas actuellement d'outils statistiques alternatifs, de type Altmetrics, pour mesurer la consultation en ligne ou le téléchargement de chaque article, bien que la fréquentation générale du site soit mesurée par des compteurs de visite.

La revue est présente dans les bases d'indexation les plus usitées en sciences biologiques : BIOSIS Previews, CAB Abstracts, MEDLINE, Scopus, ISI Web of Science...

### **2.2.3 Le Service de l'administration, de la logistique et des publications de l'OIE**

Le siège parisien de l'Organisation regroupe quatre-vingt-dix-huit agents (aux statuts divers, puisque certains membres du personnel sont détachés/mis à disposition par leur administration). Le Service de l'administration, de la logistique et des publications, placé sous la direction du Dr Daniel Chaisemartin, regroupe l'essentiel des fonctions support de l'OIE. Ce service a un effectif permanent de 22 personnes.

**Figure 1 : Organigramme simplifié du siège de l'OIE (source : OIE)**



Je détaille ci-après trois activités du service en lien avec le projet : les systèmes d'information, l'édition et la documentation.

### 2.2.3.1 Les systèmes d'information

Le pilotage des systèmes d'information (qui inclut la supervision des moyens informatiques des cinq délégations régionales de l'OIE) est placé sous la responsabilité de Bertrand Flahault, Chef de la Cellule systèmes de gestion et événements, et 1er Adjoint au chef de service, et de Daniel Chaisemartin, chef du Service.

Il est entouré d'une chef de projet informatique, d'une ingénieur informatique et d'un webmaster-développeur-support utilisateur.

### **2.2.3.2 L'édition**

L'activité éditoriale est regroupée au sein de la cellule des publications, sous la responsabilité d'Annie Souyri et du chef de service. Elle est entourée d'une secrétaire bilingue, d'une secrétaire de rédaction, d'une secrétaire traductrice, d'une gestionnaire de ventes et marketing et d'une graphiste (partagée avec l'unité « communication »). Il est régulièrement fait appel à des collaborateurs externes (traducteurs, relecteurs et réviseurs chargés de la préparation éditoriale des articles acceptés).

Les publications de l'OIE regroupent des ouvrages périodiques (le *Bulletin* [sur l'actualité de l'OIE], la *Revue scientifique et technique*, les normes internationales [code et manuel relatifs aux animaux terrestres, code et manuel relatifs aux animaux aquatiques]), des séries techniques, des ouvrages thématiques, des actes et recueils. Les bulletins, codes et séries techniques sont disponibles en trois éditions (une pour chacune des langues officielles de l'OIE<sup>2</sup>). La *Revue scientifique et technique* est trilingue et les autres publications sont généralement en anglais et peuvent être traduites selon le cas en français et/ou espagnol.

---

<sup>2</sup> Les trois langues officielles de l'OIE sont le français, l'anglais et l'espagnol



### **2.2.3.3 La documentation**

Depuis 1990, l'OIE dispose également d'un centre de documentation de grand intérêt scientifique. Aline Rousier, diplômée du cursus Documentation de l'IEP Paris, occupe à temps partiel le poste de documentaliste depuis octobre 2013, remplaçant la précédente documentaliste partie à la retraite.

Le centre de documentation vise notamment à valoriser la mémoire de l'organisation en organisant une gouvernance documentaire. Pour cela, Aline Rousier s'appuie sur un thésaurus bilingue interne de 6 000 termes spécialisés, un plan de classement des documents imprimés et numériques, des opérations de numérisation des archives de l'OIE. Elle produit également plusieurs veilles thématiques et effectue des recherches documentaires ponctuelles.

#### *2.2.3.3.1 La base documentaire*

La base documentaire a été créée en 1993 pour les besoins de l'organisation et la partie accessible au public a été mise en ligne sur le site Web de l'OIE en 2011. Avec plus de 2,5 millions de requêtes en 2014, elle est l'un des espaces le plus consulté du site Web de l'OIE.

Alimentée par la documentaliste à partir de sources internes et externes, la base documentaire répertorie les publications de l'OIE depuis sa création, ainsi que la littérature scientifique utile aux agents :

- l'ensemble des publications périodiques et ponctuelles de l'OIE depuis sa création (*Bulletin* de l'OIE, *Revue scientifique et technique*, codes et manuels, ouvrages thématiques, etc.),

- les rapports des réunions des Commissions régionales, Commissions spécialisées, groupes ad hoc et de travail,
- les résolutions, recommandations, thèmes techniques et rapports finaux des Sessions générales,
- les actes des conférences mondiales et régionales de l'OIE,
- les principales présentations powerpoint des agents effectuées lors de réunions, ateliers, séminaires, conférences,
- les brochures, fiches techniques, rapports annuels et documents de communication de l'OIE,
- les notices des ouvrages de la bibliothèque de l'OIE (accès restreint aux agents OIE),
- une sélection d'articles et rapports externes relatifs aux centres d'intérêt de l'OIE (accès restreint aux agents OIE).

La base contient plus de 15000 enregistrements, dont environ 6500 sont ouverts au public.

Plus de la moitié des documents sont disponibles en texte intégral dans les trois langues de l'OIE (français, anglais, espagnol).

#### 2.2.3.3.2 *Le logiciel Alexandrie*

Au cours de l'été 2015, Aline Rousier finalisait la migration de la base documentaire vers la nouvelle version (V7) du logiciel de gestion documentaire Alexandrie édité par la société française GB Concept. Il s'agit d'une plateforme documentaire écrite en langage Java EE, hébergée par l'OIE sur un serveur d'application Glassfish, avec des données stockées dans une base MySQL. Le logiciel est commercialisé sous la forme d'une licence annuelle, avec d'éventuels coûts annexes de formation ou d'intégration.

Alexandrie permet une gestion des fonds physiques et numériques, avec de nombreuses fonctions bibliothéconomiques associées, des outils de veille...

Accessible depuis un navigateur, cette version dite *full web* du logiciel permet de générer une interface de type extranet. Le lancement de ce nouveau portail internet trilingue est une priorité. Sur la partie publique, les versions numériques de certaines publications de l'OIE sont librement consultables. Le cas échéant, un lien propose la redirection vers la partie marchande du site pour acquérir la version imprimée. Des documents et des fonctionnalités complémentaires sont offerts aux utilisateurs enregistrés, avec un accès personnalisé et sécurisé grâce à un identifiant et un mot de passe.

Si toutes les fonctionnalités du logiciel n'ont pas vocation à être exploitées à court terme, elles permettent de préparer l'activité documentaire à des défis futurs. Par exemple, dans l'optique du web sémantique (appelé aussi web de données), il est déjà possible d'exposer le thésaurus au format RDF [1, BERMES, p. 39].

Parce qu'Alexandrie facilite la gestion des notices documentaires (formulaires de saisie, traitements par lots...), permet l'export de données structurées (notamment l'export de données exprimées en Dublin Core dans des fichiers aux formats .txt ou .xml), et inclut dans sa dernière version (V7) un module OAI-PMH, l'exposition des métadonnées peut être repensée dès à présent. C'est le projet qui m'a été confié.

## 2.3 Les objectifs de visibilité : priorités

La migration de la base documentaire de l'OIE permettra une transformation et une diffusion plus faciles des métadonnées des publications de l'OIE, pour un coût marginal négligeable. La visibilité des articles de la *Revue scientifique et technique* peut être repensée à partir de ce nouvel outil.






















En raison du regroupement de grands acteurs et de l'émergence de pratiques normatives pour la diffusion des métadonnées, il apparaît aujourd'hui que des actions modestes peuvent avoir un impact fort sur la qualité du référencement.

Plusieurs leviers nouveaux sont donc disponibles pour garantir une visibilité optimale de la revue. Ils sont présentés ici à partir de quatre grandes stratégies de recherche d'information. En effet, il s'agit bien de partir des usages des lecteurs potentiels pour aboutir à des recommandations sur les actions à conduire. Cette distinction par stratégie de recherche reste cependant artificielle puisque les services étudiés exploitent et croisent les ressources des uns et des autres.

Afin de comprendre l'étendue des sources d'information numériques disponibles en sciences vétérinaires, j'ai exploité un tutoriel recommandé par l'enseignement supérieur britannique [17, JISC]. Bien que destiné aux étudiants du domaine, ce document m'a permis de mieux identifier les étapes de la recherche d'information en sciences vétérinaires. Mais parce que la *Revue scientifique et technique* s'adresse principalement aux praticiens et décideurs, je me suis également aidé d'un article plus général, présentant les navigations des lecteurs de revues académiques pour trouver les articles pertinents pour leur recherche [7, INGER]. Bien que la grande variété des pratiques de recherche d'information et de lecture selon les disciplines [4, WARE, pp. 57-58] ne permette pas de pondérer l'usage de chaque stratégie de recherche, j'ai identifié quatre grandes navigations possibles, détaillées ci-après. En parallèle, je

présente les six leviers d'action recommandés et leurs impacts sur chaque navigation. La mise en œuvre de ces leviers est l'objet de la troisième partie de ce mémoire.

**Tableau 2 : Les quatre grandes navigations possibles pour la recherche d'information en ligne, en sciences vétérinaires et en sciences biologiques, et les leviers d'action pour y accroître la visibilité de la *Revue scientifique et technique* de l'OIE**

Stratégie de recherche d'information	Types de services utilisés	Leviers d'action
 <ul style="list-style-type: none"> <li>Moteurs de recherche</li> </ul>	 	<ul style="list-style-type: none"> <li>Optimisation du référencement naturel (SEO)</li> <li>Identifiants pérennes (DOI - CrossRef) + PMID</li> </ul>
 <ul style="list-style-type: none"> <li>Bases d'indexation</li> </ul>	    	<ul style="list-style-type: none"> <li>Identifiants pérennes (DOI - CrossRef) + PMID</li> <li>Positionnement sur l'Open access</li> <li>Exports XML</li> </ul>
 <ul style="list-style-type: none"> <li>Catalogues de bibliothèques</li> <li>Bases de connaissance (outils de découverte, résolveurs de liens...)</li> </ul>	     	<ul style="list-style-type: none"> <li>Optimisation du référencement naturel (SEO)</li> <li>Identifiants pérennes (DOI - CrossRef) + PMID</li> <li>Positionnement sur l'Open access</li> <li>Exports XML</li> <li>Serveur OAI PMH</li> <li>Fichiers KBart</li> </ul>
 <ul style="list-style-type: none"> <li>Citations</li> <li>Réseaux sociaux professionnels</li> </ul>	   	<ul style="list-style-type: none"> <li>Optimisation du référencement naturel (SEO)</li> <li>Identifiants pérennes (DOI - CrossRef) + PMID</li> <li>Positionnement sur l'Open access</li> </ul>

### 2.3.1 Les moteurs de recherches

Les moteurs de recherche grand public, au premier rang desquels se situe Google – et toutes ses versions régionales –, sont souvent la première étape des recherches d'information. Le service Google Scholar, populaire auprès des chercheurs et des praticiens, propose une recherche limitée aux publications académiques. La pertinence des résultats proposés par Google Scholar est particulièrement reconnue, y compris

dans les sciences biologiques, et supporte bien la comparaison avec de nombreux outils de découverte mis à disposition dans les bibliothèques universitaires [6, CICCONE]. Google et Google Scholar constituent donc des objectifs primordiaux en terme de visibilité.

Lorsque j'ai testé une recherche sur [www.google.com](http://www.google.com) pour un article du numéro 33 (3) de la *Revue scientifique et technique* (« Animal trypanosomosis: making quality control of trypanocidal drugs possible ») :

- la première réponse pointait vers le fichier PDF de l'article, dans sa version acceptée pour publication mais non dans sa version publiée ;
- la deuxième réponse pointait vers la notice PubMed de l'article. Malheureusement, une fois sur PubMed, aucun lien vers le texte intégral ne peut être proposé ;
- dans les réponses suivantes, aucun lien n'était proposé vers l'article sur le site de l'OIE.

La même recherche effectuée depuis Google Scholar donnait un lien vers un service affilié à PubMed. Il s'agissait d'une reprise de la notice du site américain, toujours sans lien vers le texte intégral ni même une référence au site de l'OIE.

Afin de garantir la consultation des articles sur le portail documentaire de l'OIE, j'ai vérifié l'absence d'appropriation intempestive des fichiers PDF par Google Scholar.

À l'inverse de Google Books (où les livres – numérisés à partir des fonds de bibliothèques partenaires ou transmis par les éditeurs – sont consultés depuis les serveurs de Google), Google Scholar ne propose aujourd'hui que des liens externes vers le texte intégral des articles. Deux cas de figure :

- le texte intégral a été indexé par Google à partir du site de l'éditeur : Google Scholar propose un lien direct vers l'article sur le site de l'éditeur.
- le texte intégral a été indexé par Google à partir d'un site autre que celui de l'éditeur : Google Scholar propose un lien vers une notice de l'article et, en vis-à-vis, un lien vers le texte intégral, en précisant la source.

Conformément aux recommandations de Google Scholar aux éditeurs (< <https://scholar.google.com/intl/en/scholar/inclusion.html> >, consulté le 15 septembre 2015), mon travail a donc porté sur une optimisation du référencement naturel (SEO ou *Search engine optimization*) des pages du portail documentaire de l'OIE, et sur la mise en place d'identifiants pérennes.

### **2.3.2 Les bases d'indexation**

En enrichissant et en agrégeant les données pour toutes les publications d'un même champ scientifique, les bases d'indexation (appelées aussi services A&I pour *abstracting and indexing services*) apportent une valeur ajoutée précieuse pour toute recherche d'information. Concrètement, ces bases recensent les articles d'un ensemble de publications périodiques, en compilant les résumés (*abstracts*) et attribuant des mots-clés issus de vocabulaires contrôlés (thésaurus, nomenclatures ou ontologies spécialisés).

D'autre part, les données présentes dans ces bases servent parfois au calcul d'indices bibliométriques (dont l'Impact factor de JCR). Ces mêmes indices peuvent d'ailleurs y être utilisés pour trier l'affichage des résultats de recherche.

Il m'est donc apparu incontournable de faire un point sur les bases qui indexent les articles de la *Revue scientifique et technique*. Pour cela, je me suis appuyé sur les

informations présentes sur le site internet de l'OIE, dans les pages intérieures de la revue et dans le catalogue des publications. D'un avis général à l'OIE, ces informations nécessitaient d'être revalidées et actualisées si nécessaire.

Je présente ci-après les principales bases, ainsi que les leviers d'action pour améliorer la présence de la revue (l'objectif étant un référencement exhaustif, au plus tôt de la date de parution, avec un lien vers la notice de l'article sur le portail documentaire de l'OIE où le fichier PDF du texte intégral est disponible en libre accès).

### **2.3.2.1 MEDLINE/PubMed – NCBI**

PubMed < <http://www.ncbi.nlm.nih.gov/pubmed> > est une ressource gratuite développée et maintenue par le National Center for Biotechnology Information (NCBI), à la Bibliothèque nationale de médecine des États-Unis (US National Library of Medicine, NLM). Ses 24 millions de références en littérature biomédicale sont essentiellement issues de la base d'indexation MEDLINE (Medical Literature Analysis and Retrieval System Online) produite par la Bibliothèque nationale de médecine des États-Unis. La qualité de l'indexation dans MEDLINE repose sur un travail d'analyse approfondi, effectué par des professionnels. Le thésaurus utilisé est MeSH (Medical Subject Headings), exploité par une majorité d'outils de recherche et de veille documentaire.

PubMed couvre la biomédecine et la santé, et certains thèmes en sciences de la vie, sciences du comportement, chimie et bio-ingénierie.

Depuis 1990, 1 832 articles de la *Revue scientifique et technique* ont été indexés par MEDLINE, avec leurs résumés (*abstracts*). Les notices, consultables dans PubMed, ont



chacune un identifiant unique appelé PMID (pour PubMed ID). Ces PMID sont précieux car fréquemment utilisés comme identifiants uniques par les outils de citation.

L'indexation de chaque numéro de la *Revue scientifique et technique* est effective après un délai assez long de 4 à 6 mois.

En l'absence d'identifiant pérenne de type DOI, aucun lien hypertexte vers le site de l'OIE n'est actuellement intégré dans les notices MEDLINE. Ainsi, lorsque ces notices sont versées dans PubMed, le site ne mentionne pas l'OIE et propose un lien externe pointant vers le résolveur de lien de la société EBSCO. Faute de métadonnées appropriées, ce résolveur indique – à tort – qu'aucune source n'est disponible pour le texte intégral.

Parce que PubMed est un service très populaire dans le monde entier pour les recherches documentaires en sciences biologiques, mes recommandations ont visé :

- l'accélération du signalement des nouveaux articles
- la présence dans chaque notice PubMed d'un lien hypertexte vers la notice de l'article correspondant dans la base documentaire de l'OIE
- la présence dans chaque notice de la base documentaire de l'OIE de la référence PMID correspondante

Pour cela, il s'agira notamment de fournir à PubMed des fichiers XML conformes à leur schéma de métadonnées (.xsd) afin d'obtenir un premier signalement minimal des articles sur le site, dans l'attente de l'indexation par MEDLINE. Ceci permettra d'être labellisé PubMed citation data provider, préalable obligatoire à la fourniture de liens hypertexte au travers du programme LinkOut.

Grâce à une extraction de fichier proposée depuis PubMed, j'ai pu aligner les notices de la base documentaire de l'OIE avec les références PMID correspondantes. Ce fichier des équivalences strictes entre Alexandrie et PubMed permettra d'enrichir rétrospectivement PubMed, via le programme LinkOut, avec les identifiants pérennes (DOI) de chaque article. Comme indiqué par l'équipe en charge des relations éditeurs (PubMed Data Provider Support Team), un simple fichier .txt sera nécessaire, listant sur deux colonnes l'alignement entre la référence OIE et son PMID.

### **2.3.2.2 CAB Abstracts – CABI**

La base CAB Abstracts est produite par l'organisation internationale CABI (Commonwealth Agricultural Bureaux International), dont le siège est en Grande-Bretagne. Elle référence plus de 7,9 millions de documents, de 1973 à ce jour, avec plus de 360 000 nouvelles entrées par an. Elle couvre les sciences du vivant appliquées, telles que les sciences de l'agriculture et de l'alimentation, la nutrition, l'environnement, les sciences vétérinaires, l'économie.

CAB Abstracts est commercialisée sous différentes formes et notamment au travers de sous-ensembles thématiques. La *Revue scientifique et technique* est notamment présente dans les bouquets :

- Animal Science,
- Veterinary Bulletin,
- Index Veterinarius (dont l'actualisation est abandonnée).

La revue y est actuellement indexée à partir de l'édition imprimée et les notices des articles ne contiennent pas de lien hypertexte vers le texte intégral disponible sur le site de l'OIE. Les bonnes relations entre les deux organisations ont permis de prendre contact avec CABI. À leur demande, il est envisagé de fournir un fichier PDF unique pour chaque nouveau numéro de la revue et d'organiser un flux de données au format XML.

### **2.3.2.3 Web of Knowledge – Thomson Reuters**

La base Web of Knowledge est produite par Thomson Reuters et ses filiales. Le sous-ensemble commercialisé sous le nom Web of Science regroupe les bouquets suivants (eux-même commercialisés indépendamment), dans lesquels la *Revue scientifique et technique* est référencée :

- **BioSciences Information Service**
  - Biological Abstracts
  - BIOSIS Previews
  
- **Current Contents**
  - Agriculture, Biology & Environmental Sciences
  - Agriculture/Agronomy
  - Agricultural Chemistry
  - Animal Sciences

- Aquatic Sciences
  - Biology
  - Biotechnology & Applied Microbiology
  - Entomology/Pest Control
  - Environment/Ecology
  - Food Science/Nutrition
  - Multidisciplinary
  - Plant Sciences
  - Veterinary Medicine/Animal Health
- 
- **Food science and Technology Abstract**
- 
- **Journal Citation Reports**
  - Science Citation Index
  - SciSearch
- 
- **Zoological Record**

Actuellement, tous les articles de la *Revue scientifique et technique* sont indexés « revue en main », sans lien hypertexte vers le site de l'OIE. Pour d'évidentes raisons

d'efficacité dans l'intégration des abstracts, mon interlocutrice chez Thomson Reuters (service des relations éditeurs) souhaitait obtenir par transfert FTP un fichier PDF (non verrouillé et regroupant l'ensemble des articles) pour chaque nouveau numéro de la revue. Afin d'encadrer les usages d'un tel fichier, des aspects contractuels restaient à évaluer.

Dans un deuxième temps, il conviendra d'adresser des fichiers XML. Mais cette étape devra attendre le signal de Thomson Reuters : mon interlocutrice indiquait qu'en l'absence de schéma XML prédéfini, son service avait momentanément un nombre élevé de tests et de cartographies de données en souffrance pour les fichiers XML fournis par les éditeurs partenaires.

#### **2.3.2.4 SCOPUS – Elsevier**

Principal concurrent de Web of Knowledge (Thomson Reuters), Scopus est la base d'indexation de Elsevier. Scopus propose également des outils d'analyse des citations d'une personne, d'un groupe, d'un article ou d'un journal. Ce service référence 21 000 journaux scientifiques (y compris 1 200 titres en libre accès), 600 publications industrielles et 350 collections d'ouvrages.

La *Revue scientifique et technique* est indexée dans Scopus (à partir de l'édition imprimée, comme pour Thomson Reuters). Mais étrangement la revue n'y est pas listée dans le bouquet « Veterinary ». D'autre part, les notices n'intègrent pas de lien hypertexte vers le site de l'OIE.

Il convient donc de maintenir l'envoi de l'édition imprimée et de contacter Scopus pour proposer les métadonnées des articles afin d'accélérer et de compléter le référencement.

### **2.3.2.5 AGRICOLA – US National Agricultural Library**

AGRICOLA (AGRICultural OnLine Access) < <http://agricola.nal.usda.gov> > est le catalogue de la National Agricultural Library (États-Unis). Cette ressource est très utilisée en Amérique du Nord, en raison de sa spécialisation, complémentaire à PubMed.

Tous les articles de la *Revue scientifique et technique* y sont indexés depuis 1982 (contrairement à 1990 pour MEDLINE).

Afin d'accélérer et d'améliorer le processus d'indexation, Agricola invite les éditeurs à leur fournir des fichiers XML par transfert FTP. Le schéma (.xsd) exigé est celui de PubMed. En conséquence, dès que les fichiers de métadonnées seront validés par PubMed, il sera envisageable d'alimenter également la base Agricola, sans développement particulier.

### 2.3.2.6 DOAJ

Le Directory of Open Access Journals < [www.doaj.org](http://www.doaj.org) > est un annuaire des revues académiques disponibles en libre accès. Les revues référencées peuvent y déclarer tous leurs articles (à condition de les auto-indexer). En conséquence, l'entrepôt OAI-PMH du DOAJ est moissonné par de nombreuses grandes bibliothèques et bases de connaissance afin d'enrichir leurs catalogues avec des notices de contenus librement accessibles.

Au 1er août 2015, l'annuaire contenait :

- 10 491 revues
- 6 354 revues avec déclaration des articles
- 134 pays
- 1 939 916 articles

DOAJ représente en quelque sorte les « puristes » du libre accès, autour du modèle dit de la voie dorée (voir infra 1.2.2. Le libre accès). En conséquence, les conditions de référencement dans l'annuaire sont aujourd'hui très restrictives : seules sont éligibles les revues conformes à certains critères de qualité et fermement engagées dans le libre accès.

Il semble incontestable qu'un référencement dans DOAJ apporterait une visibilité forte aux articles de la *Revue scientifique et technique*, qui n'y sont actuellement pas présents. Comme le DOAJ est entièrement gratuit, le processus d'inscription est très long (4 à 6 mois) et surtout très contraignant pour les éditeurs. Sans anticiper sur les orientations futures de l'OIE, j'ai étudié en détail chaque critère d'éligibilité. Cette analyse, dont certains éléments sont présentés dans la troisième partie de ce

mémoire, permet en effet un audit rapide du positionnement de l'organisation sur le libre accès.

Enfin, notons que dans l'éventualité où l'OIE serait candidate au référencement dans DOAJ, il conviendrait de créer un fichier XSLT pour obtenir des fichiers XML conformes au schéma de métadonnées (.xsd) exigé. Ce schéma permet en effet d'unifier les métadonnées exposées sur l'entrepôt OAI-PMH de DOAJ.

### **2.3.2.7 AGRIS – FAO**

La base Agris (International System for Agricultural Science and Technology) < <http://agris.fao.org> > est produite par la FAO. Multilingue et en libre accès, elle contient plus de 7 millions de références bibliographiques en sciences et technologie agricoles, alimentation, environnement, etc. Les articles sont indexés à partir du thésaurus AGROVOC de la FAO. Agris agrège les données de plusieurs autres sources (Europeana, Banque mondiale, DBPedia...).

Après vérification, la *Revue scientifique et technique* n'est plus référencée dans Agris depuis 2010. En effet, une requête sur « OIE » et l'ISSN de la revue (0253-1933) remonte 1284 références mais seules 6 sont postérieures à 2005 et il n'y en a plus depuis 2010. Des requêtes plus ouvertes, ou avec des titres d'articles, ont également confirmé l'arrêt du référencement.

J'ai donc recommandé une prise de contact avec la FAO et ses partenaires pour relancer le référencement dans Agris. Conformément aux instructions données aux éditeurs, il conviendra ensuite de :

- 1/ S'enregistrer comme fournisseur de métadonnées



2/ Renseigner les informations sur les collections de l'OIE

3/ Transmettre les métadonnées, au choix :

- En pointant vers le serveur OAI-PMH de l'OIE (données exprimées en Dublin Core)
- En transmettant à intervalles réguliers des fichiers XML conformes au schéma (.xsd) PubMed

#### **2.3.2.8 HINARI – OMS**

Le programme HINARI < <http://extranet.who.int/hinari> > n'est pas à proprement parler une base d'indexation mais je le mentionne ici puisqu'il est utilisé dans certains pays en voie de développement comme point d'entrée aux ressources numériques. Il permet d'accéder à de vastes fonds éditoriaux en biomédecine et en santé.

Les éditeurs partenaires de l'OMS s'engagent à pratiquer des tarifs adaptés et soutenables. Dans les autres cas, HINARI s'appuie sur le service DOAJ pour signaler les ressources en libre accès, sans préciser leur contenu. En conséquence, la base ne propose pas d'index de recherche d'articles ni de résumés.

Les publications de l'OIE n'étaient pas référencées dans Hinari. Suite à une demande d'information à l'OMS sur les conditions de référencement, la *Revue scientifique et technique* et le *Bulletin* y sont désormais signalés depuis le 20 juillet 2015.

### **2.3.2.9 The Capsule Report**

The Capsule Report est un panorama de presse mensuel payant, à destination des vétérinaires nord-américains. Depuis 1981, cette publication signale et classe les articles, livres, conférences ou actes de congrès jugés pertinents pour son lectorat.

Bien que The Capsule Report soit utilisable comme outil de recherche, en raison de son ancienneté et des thématiques suivies, son absence d'exhaustivité limite la stratégie de référencement de l'OIE.

En conséquence, j'ai recommandé d'agir en priorité sur la veille utilisée par les rédacteurs. Par exemple, une veille sur une entrée de la nomenclature MeSH (Medical Subject Headings, maintenue par la Bibliothèque nationale de médecine des États-Unis) doit détecter les articles de la *Revue scientifique et technique* sur le sujet. Pour cela, il convient de garantir le meilleur référencement possible dans les bases MEDLINE-PubMed et Agricola, qui toutes deux indexent avec la nomenclature MeSH et font partie des sources utilisées par The Capsule Report.

### **2.3.2.10 Bases abandonnées**

#### *2.3.2.10.1 Focus On: Veterinary Science and Medicine*

FoVS&M (Focus On: Veterinary Science and Medicine), base spécialisée autrefois présente sur les plateformes Ovid (Wolters Kluwer) et Web of Knowledge (Thomson Reuters), n'est plus commercialisée par ces sociétés depuis plusieurs années. À ce jour, il n'a pas été possible de contacter l'éditeur de la base pour une confirmation de son arrêt.

#### 2.3.2.10.2 *Index Medicus*

Parfois encore mentionné comme publication référençant la *Revue scientifique et technique*, l'Index Medicus (IM) s'est arrêté en 2004 avec le volume numéro 45, lorsque la base de données MEDLINE a pris son relais. Publié depuis 1879 par la Bibliothèque nationale de médecine des États-Unis (US National Library of Medicine, NLM), il compilait les tables des matières des principales revues biomédicales.

#### 2.3.2.10.3 *Littérature vétérinaire francophone*

Éditée par la faculté de médecine de Montréal, cette base recensait depuis 1985 des articles de revues, conférences, thèses, chapitres de livres... Elle semble abandonnée depuis plusieurs années (dernière trace en 2005). À ce jour, il n'a pas été possible de contacter l'éditeur de la base pour une confirmation de son arrêt.

#### 2.3.2.10.4 *PASCAL – INIST*

PASCAL était une base d'indexation alimentée par l'INIST-CNRS (Institut national pour l'information scientifique et technique), arrêtée le 31 décembre 2014. Les archives seront ouvertes en libre accès au cours de l'année 2015.

Il est utile de noter que, prenant acte de l'évolution des besoins des chercheurs du CNRS, l'Inist-CNRS a abandonné l'indexation des bases PASCAL et FRANCIS et utilise actuellement un outil de découverte (service « Discovery tool » de la société EBSCO) que je présenterai ci-après.

### **2.3.3 Les catalogues de bibliothèques et les bases de connaissance**

En raison de la complexité actuelle des collections de périodiques numériques et des droits de consultation associés, les grandes bibliothèques utilisent des bases de connaissance pour gérer leurs collections numériques.

Ces bases utilisent notamment la norme OpenURL (*ANSI/NISO Z39.88-2004, The Open URL Framework for Context-Sensitive Services*, disponible à < [http://www.niso.org/apps/group\\_public/project/details.php?project\\_id=82](http://www.niso.org/apps/group_public/project/details.php?project_id=82) >, consulté le 15 septembre 2015) : il s'agit d'adresses URL qui embarquent un certain nombre de métadonnées contextuelles (telles que le genre, l'auteur, les mots du titre, l'ISSN, l'année...), souvent contenues dans une référence bibliographique issue d'une base d'indexation. Les métadonnées transportées dans l'OpenURL sont analysées par des logiciels commerciaux appelés résolveurs de liens, qui construisent une requête vers la cible appropriée. La cible est généralement un document référencé au sein de la base de connaissance, avec les droits de consultation associés.

Concrètement, un tel dispositif permet à un utilisateur qui identifie un lien hypertexte vers un article intéressant d'accéder « automatiquement » au document, conformément aux droits de consultation associés à son profil (par exemple droit d'accès au texte intégral s'il s'agit d'une ressource payante). En résumé, la base de connaissance d'une bibliothèque contient les données relatives aux ressources détenues et les syntaxes de liens URL permettant d'accéder aux cibles externes telles

les revues auxquelles la bibliothèque est abonnée et qui sont hébergées sur les serveurs informatiques des éditeurs.

Cet écosystème est complété depuis quelques années par les outils de découverte, interfaces d'interrogation uniques sur de vastes ensembles de ressources disponibles. Les quatre principaux outils du marché sont Summon de ProQuest, Primo de ExLibris, Discovery Service de Ebsco et Discovery Services de OCLC. Si la popularité et l'efficacité de ces « métamoteurs » pour la recherche d'information n'est pas contestée, il existe peu de données disponibles pour en mesurer les effets concrets en terme de consultation des revues [8, LEVINE-CLARK].

Entre autres sujets, une critique récurrente sur la relative opacité des résultats proposés par les outils de découverte a conduit les principaux acteurs à initier l'Open Discovery Initiative (ODI), groupe de travail réuni au sein du National Information Standards Organization (NISO, organisme américain à but non lucratif dédié aux normes dans le domaine de l'édition, des bibliothèques et de l'accès à l'information). Un livre blanc sur le futur des outils de découverte [5, BREEDING] a été publié qui éclaire les enjeux et problématiques à venir, notamment sous l'angle des relations entre les sociétés qui produisent ces outils et les éditeurs, et sur les données liées.

À ce tableau, il convient d'ajouter les projets de base de connaissance libres et ouvertes lancés ces dernières années dans l'enseignement supérieur français (projet BACON) et britannique (KB+ à vocation locale et GoKB à vocation internationale). En s'affranchissant des produits commerciaux, ces projets visent à fournir des informations sur les ressources électroniques disponibles sur le marché académique. Puisqu'il s'agit de données ouvertes, ils pourraient bénéficier à une communauté plus large que les seules bibliothèques universitaires. Chacune de ces bases rassemblera les informations de publication essentielles :

- données bibliographiques (titre, bouquet, plateforme...) ;

- informations d'abonnement (dates de renouvellement, périodes de préavis, etc.) ;
- informations sur les licences (conditions d'accès distant, utilisateurs invités, accès post-désabonnement, etc.).

Le projet de BAs de COnaissances Nationale (BACON), porté par l'ABES (Agence bibliographique de l'enseignement supérieur), mettra à disposition les données concernant les bouquets de périodiques et de livres électroniques proposés aux établissements français d'enseignement supérieur et de recherche.

Knowledge Base Plus (KB+ < <http://www.kbplus.ac.uk> >), qui émane de l'organisme public britannique JISC (Joint Information Systems Committee), opérera à l'échelle nationale de la Grande-Bretagne, au niveau consortial et sera le réceptacle des informations sur les licences nationales.

Global Open Knowledge base (GOKb < <http://gokb.org> >) est une collaboration multi-institutionnelle qui comprend des partenaires des États-Unis et en Europe. Le projet de GOKb est conçu et mis en œuvre par les partenaires fondateurs Quali OLE et JISC. GOKb fonctionnera à l'échelle internationale et contiendra les informations sur les éditeurs, sur les bouquets, les licences standards. Les données seront mises gratuitement à la disposition des bibliothèques, éditeurs universitaires, prestataires de services de la bibliothèque, et du public à l'aide d'une interface de programmation (API, *Application Programming Interface*).

BACON est partenaire de KB+ et de GOKb : les informations communiquées par les éditeurs francophones seront reprises dans les trois bases, améliorant ainsi la visibilité des publications.

Au vu de ce qui précède, il s'avère essentiel que les métadonnées de la *Revue scientifique et technique* soient présentes dans les bases de connaissance et outils de découverte, et exploitables par les résolveurs de liens. Or, bien que la revue soit présente dans plusieurs catalogues de grandes bibliothèques, elle est souvent signalée uniquement comme document imprimé, avec une notice descriptive généralement limitée aux numéros, sans descendre jusqu'aux articles.

Plusieurs bases de connaissance et outils de découvertes reprennent les indexations de PubMed ou d'autres bases d'indexation, et signalent donc les articles de la revue. Cependant, d'une part, ce référencement intervient avec un décalage important par rapport à la date de mise en ligne de l'article. D'autre part, en l'absence d'identifiants pérennes et de métadonnées adaptées, les résolveurs de liens associés aux outils de découverte ne peuvent pas « reconstruire » l'URL vers la page de l'article sur le site de l'OIE.

Nous verrons plus loin qu'en passant par une agence de DOI, l'OIE pourra contourner ces difficultés. De plus, les recommandations NISO aux éditeurs pour la fourniture de métadonnées aux bases de connaissance définissent le protocole OAI-PMH et les fichiers KBART comme les deux éléments complémentaires indispensables. Je présenterai dans la troisième partie de ce mémoire les actions conduites en ce sens.

Notons cependant que les bases de connaissance, notamment BACON, exigent que les données descriptives déposées sur les entrepôts OAI-PMH et celles transmises au travers des fichiers KBART soient placées sous une licence Creative Commons de type CC0 (< <https://creativecommons.org/publicdomain/zero/1.0/deed.fr> >). Cette licence permet l'exposition et la transformation maximale des métadonnées, y compris pour un usage commercial, sans nécessité de citer le producteur des données. Sur cette question, il pourra être utile de consulter les réflexions en cours dans l'univers des données ouvertes (*open data*) [par exemple 1, BERMES, pp. 56-59 et 71-74].

### **2.3.4 Les outils de citations et les réseaux sociaux académiques, scientifiques et professionnels**

Qu'il s'agisse d'une recommandation par un tiers ou d'une référence en fin d'article, de nombreux documents sont consultés parce qu'ils ont été cités dans un contexte qui validait leur pertinence. Pour garantir la qualité des informations échangées et favoriser ces recommandations, il était utile de se pencher sur les outils de citations (utilisés par les chercheurs pour gérer leurs références bibliographiques) et sur les réseaux sociaux académiques, scientifiques et professionnels (nouveaux lieux d'échange de références bibliographiques).

#### **2.3.4.1 Les outils de citations**

Plusieurs logiciels permettent d'importer et d'exporter les références d'articles afin de pouvoir les citer selon un format approprié.

Qu'ils soient commerciaux (EndNote de Thomson Reuters, RefWorks de ProQuest, Mendeley de Elsevier, Papers de Springer) ou libres et open source (Zotero), ces outils proposent de récupérer les métadonnées des articles :

- par extraction des données présentes dans les balises méta de la page HTML de l'article ;
- par importation d'un fichier (majoritairement de type RIS) téléchargé depuis la page HTML de l'article ;



- par interrogation distante de bases de connaissance, de catalogues de bibliothèques ou de CrossRef, soit à partir d'une recherche textuelle, soit à partir du DOI ou du PMID de l'article.

Une citation correctement formée favorise la visibilité d'un article. Elle facilite également la reconstruction de l'URL de l'article par les résolveurs de liens et garantit une identification univoque par les outils bibliométriques (calcul de l'Impact factor).

En l'absence de message de type « Pour citer cet article : » suivi d'une citation correctement formée, les lecteurs de la revue doivent actuellement saisir eux-mêmes les données de citation dans leur outil. D'autre part, bien que de nombreuses extensions existent pour des logiciels de gestion de contenus open source tels que Drupal, l'export de métadonnées au format RIS à partir d'Alexandrie nécessiterait des développements assez lourds (comme indiqué dans la description du format RIS disponible sur le site Thomson Reuters : < <http://kbportal.thomson.com/display/2/index.aspx?tab=browse&c=&cpc=&cid=&cat=&catURL=&r=0.583924> >, consulté le 3/8/2015).

Mes recommandations ont donc porté sur :

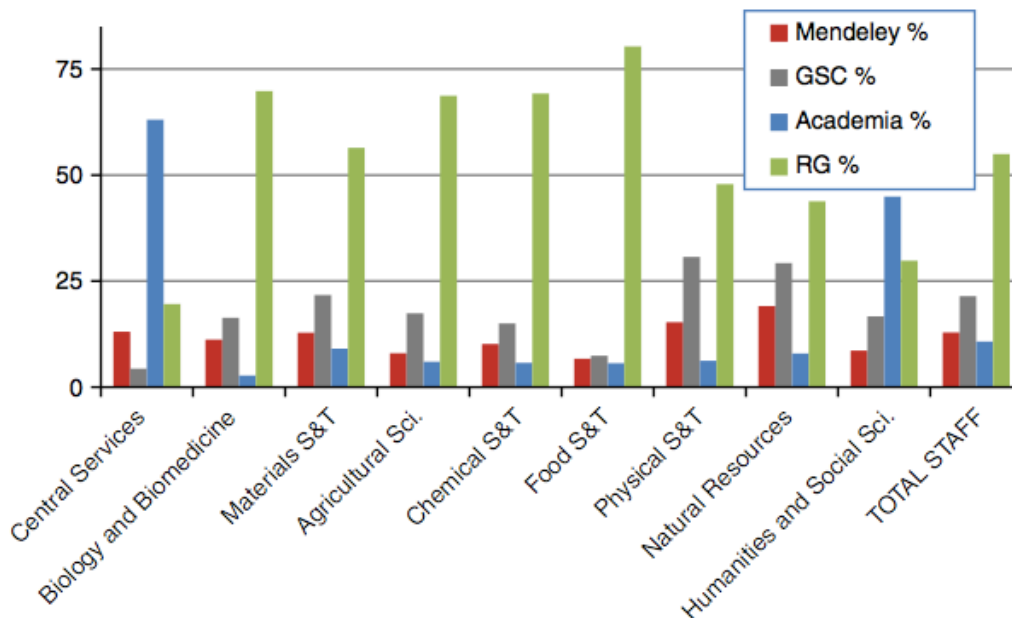
- l'extraction directe des métadonnées à partir des balises *meta* des pages HTML (scraping) ;
- l'identification univoque à partir des DOI et PMID.

### 2.3.4.2 Les réseaux sociaux académiques, scientifiques et professionnels

Les réseaux sociaux professionnels sont des espaces plébiscités par de nombreux chercheurs et praticiens pour signaler leurs propres publications et partager des références bibliographiques.

Plusieurs plateformes visent aujourd'hui à favoriser les interactions entre chercheurs. Parmi les réseaux sociaux académiques et/ou scientifiques les plus populaires, les quatre leaders sont Mendeley, Google Scholar Citations, Academia et Research Gate. La figure 2 donne une indication du poids respectif de chaque outil selon les disciplines et montre la domination de Research Gate dans les sciences dures, notamment en biologie et biomédecine.

**Figure 2 : Répartition des profils par disciplines au sein des quatre principaux réseaux sociaux académiques**



(source : ORTEGA José Luis Ortega, 2015. Disciplinary differences in the use of academic social networking sites. *Online Information Review*. 27 avril 2015. Vol. 39, n° 4, p. 525)

Academia rassemble principalement des chercheurs en sciences humaines et sociales.

Plus que des réseaux sociaux, Mendeley et Google Scholar Citations sont avant tout des outils de partages de références bibliographiques. Propriété de l'éditeur Elsevier, Mendeley encourage le partage et le réseautage entre ses membres. Le service est fondé sur des algorithmes de recommandation puissants. Il est probable que le modèle repose sur une économie de l'attention et l'étude des pratiques des communautés scientifiques, afin de proposer des produits éditoriaux ciblés.

En conséquence, mon intérêt s'est principalement porté vers Research Gate (RG) < <http://www.researchgate.net> >.

Proche par certains aspects de réseaux sociaux grand public tels que Facebook ou LinkedIn, Research Gate est un réseau social en langue anglaise destiné aux scientifiques. Lancé en 2008 par deux virologues (Dr Ijad Madisch et Dr Sören Hofmayer) et un informaticien (Horst Fickenscher), le site déclarait en mai 2015 avoir dépassé les 7 millions d'inscrits. Bill Gates figure depuis 2013 au rang des principaux investisseurs (source : < <https://www.crunchbase.com/organization/researchgate> >, consulté le 15 septembre 2015). Research Gate opère depuis Berlin et est donc soumis aux lois allemandes et européennes.

L'accès au site exige une inscription préalable, gratuite pour les scientifiques, payante pour les entreprises. Research Gate déclare ne pas revendre les données de ses utilisateurs à des tiers. Le modèle économique semble reposer sur le développement futur d'un marché payant de produits et services scientifiques. Le site fait régulièrement l'objet de critiques mettant en cause des envois massifs de mails non sollicités, une gestion défailante de l'identité numérique des personnes et des téléchargements abusifs de publications protégées.

Les scientifiques s'inscrivent principalement à RG pour contacter d'autres chercheurs ou experts, trouver des articles, promouvoir leurs recherches. Depuis leurs profils, ils peuvent suivre l'activité de leurs relations, découvrir des offres d'emploi, recevoir ou répondre à des questions relevant de leur domaine de recherche et peuvent ainsi contribuer à l'évaluation desdites questions/réponses.

Un « RG score » quantifie la réputation scientifique des inscrits à partir de toutes les contributions et interactions intervenues sur le site (pondérées par le nombre et la qualité des pairs) et de la notoriété des revues dans lesquelles la personne a publié.

Sur Research Gate, les scientifiques sont incités à déclarer leurs publications et à partager le texte intégral. Dès la phase d'inscription, le site propose d'ailleurs l'auto-complétion du profil avec les publications identifiées par le nom de l'auteur dans plusieurs bases de données (PubMed, arXiv, IEEE, CiteSeer, Cornell University library, NASA, RePEc, BioMED Central...). Research Gate se présente d'ailleurs comme un outil d'auto-archivage des publications mais, contrairement aux répertoires institutionnels, sans apporter de garantie sur la pérennité de cet archivage.

Lorsqu'un auteur souhaite ajouter sur son profil un article publié dans une revue académique, Research Gate lui propose de consulter le site Sherpa/Romeo (< <http://www.sherpa.ac.uk/romeo> > pour vérifier la politique d'auto-archivage, les droits et autorisations attachés à la revue. La personne doit en outre s'engager sur sa qualité d'auteur. Si la publication est signée par d'autres co-auteurs inscrits sur Research Gate, ceux-ci sont automatiquement notifiés par le site.

Pour les documents non publiés, et sous réserve que l'auteur soit pleinement titulaire des droits, Research Gate propose de générer un DOI pour identifier le document numérique de façon pérenne.

Notons enfin que tous les documents téléchargés dans ResearchGate sont indexés par Google Scholar.

L'un des points essentiels pour les éditeurs académiques est que, pour l'ajout d'un lien hypertexte à une citation, Research Gate impose aux auteurs le téléchargement (*upload*) du fichier PDF du texte intégral. De nombreux auteurs, soucieux d'afficher leurs publications et de répondre aux sollicitations de leur réseau, franchissent allègrement le pas malgré les droits attachés au document et le risque de dispersion des statistiques de consultation. La *Revue scientifique et technique* étant protégée par un copyright, ses auteurs peuvent citer leurs articles sur Research Gate mais sans lien externe vers le texte intégral.

Cette étape, aussi impérative que malicieuse (elle garantit au site la récupération d'un maximum de documents), pourrait bien entendu être contournée par l'attribution de DOI : l'identifiant pérenne fait partie intégrante de la citation. Conformément aux recommandations de l'agence CrossRef, le DOI peut être exprimé sous forme d'URL, sur le modèle : < [http://dx.doi.org/\[préfixeOIE\]/\[réf.article\]](http://dx.doi.org/[préfixeOIE]/[réf.article]) >.

Pour concentrer les consultations des articles de la *Revue scientifique et technique* sur le site de l'OIE, j'ai recommandé de :

- Communiquer sur les sites Sherpa/Romeo et Heloise la politique de l'OIE en matière d'auto-archivage des articles par les auteurs (quelle version du manuscrit peut être déposée, quels repositories sont autorisés, après quel délai, pour quelles exploitations dérivées, avec quels impératifs de citation ou de liens hypertexte vers le site de l'OIE, etc.) ;
- Attribuer un DOI à chaque publication de l'OIE. Les DOI font partie des citations et leur présence est donc acceptée par Research Gate. Ceci permettra aux auteurs de signaler leurs publications avec un lien vers le texte intégral en libre accès sur le site de l'OIE. Ceci évitera également les sollicitations intempestives à partir du bouton « Request Full Text ».

- Inciter les auteurs à inclure systématiquement les DOI dans les citations, par exemple en attribuant ceux-ci le plus tôt possible et en les faisant apparaître dans les documents.

Je n'ai pas proposé une animation pro-active de la page OIE sur Research Gate, d'une part parce que cette page existe déjà (générée automatiquement par le site), d'autre part parce que les données fournies par d'autres institutions ont servi à construire des campagnes de spams.

## **2.4 Conclusion intermédiaire**

Au lancement du projet, il était difficile de distinguer les besoins des solutions : un ensemble diffus de mots clés (OAI-PMH, moteurs de recherche, outils de découverte, DOI, XML, réseaux sociaux...) semblaient converger vers un même objectif de visibilité des publications. Cette étape d'analyse fondée sur la navigation de la recherche d'information a permis d'éclairer les objectifs et concepts. Ainsi, non seulement les leviers d'action ont pu être identifiés mais j'ai pu les prioriser en fonction des objectifs et du contexte propre à l'OIE.

La mise en œuvre de ces leviers d'action est présentée dans la troisième partie de ce mémoire.

### 3 Améliorer la visibilité de la *Revue scientifique et technique* de l'OIE : quelles solutions ?

---

Cette troisième partie détaille les solutions mises en œuvre pour garantir une meilleure visibilité de la *Revue scientifique et technique* lors de toute recherche d'information sur internet.

Ces solutions éprouvées correspondent à des usages et bonnes pratiques fortement répandus dans l'univers des revues académiques. Il s'agit ici d'en présenter les enjeux et d'illustrer leur implémentation dans le contexte spécifique de l'OIE ; les questions rencontrées, comme les réponses apportées, pourront intéresser et éclairer les professionnels confrontés à des problématiques similaires.

Chacune de ces solutions a fait l'objet d'une note de synthèse adressée aux équipes de l'OIE concernées. Mes préconisations furent essentiellement techniques mais j'ai porté une attention particulière à l'accompagnement du projet en matière de conduite du changement et de formation des usagers, du début du projet (analyse de l'existant, entretiens, recueil des exigences) jusqu'à sa fin (réunions de présentation et notes très descriptives sur le choix des solutions, leur fonctionnement, leurs enjeux et leurs impacts pour chacune des parties prenantes). Pour des raisons de place et confidentialité, mais surtout parce que les enjeux sont essentiellement spécifiques à l'OIE, ces préoccupations ne sont pas traitées dans le présent mémoire.

#### **3.1 Optimisation du référencement naturel**

Le référencement naturel par les moteurs de recherche (ou SEO pour *Search engine optimization*) est un ensemble de techniques qui vise à permettre aux robots

d'indexation utilisés par Google, Google Scholar ou autres, de mieux référencer le contenu d'un site internet.

Je me suis concentré sur :

- l'architecture générale du site de l'OIE, et plus précisément la navigation au sein du portail documentaire
- les instructions données aux robots d'indexation
- l'enrichissement sémantique des pages du portail documentaire

### **3.1.1 Architecture et navigation au sein du site internet de l'OIE**

L'OIE souhaite conserver la maîtrise de la chaîne éditoriale et favoriser la consultation de ses publications numériques sur son site Web (pas de recours à un agrégateur de contenus). Dans cette optique, des CMS open source, tels que OJS (Open Journal Systems) développé par PKP (Public Knowledge Project), ou commerciaux, tels que Atypon ou Highwire Press, semblent particulièrement intéressants. Ils offrent des architectures adaptées aux revues académiques, en fluidifiant le processus éditorial sans entraver le choix de tel ou tel modèle économique (libre accès, accès premium, abonnement, etc.). Une étude de ces solutions dépasserait l'objet de ce mémoire.

J'ai recherché prioritairement des solutions intermédiaires fondées sur le portail documentaire de l'OIE, impliquant des impacts modérés sur l'organisation actuelle.

La quasi-intégralité des revues académiques disposent d'un site internet dédié. Il s'agit souvent d'un simple sous-domaine du site de l'éditeur ou de la société savante. Cette



architecture garantit une indexation optimale par les moteurs de recherche et favorise la résolution de liens, en cas d'implémentation de la norme OpenURL.

Voici par exemple les informations transmises aux consortiums d'universités par Oxford University Press (OUP) au sujet de la revue *British journal of social work* :

- Title : British journal of social work
- Print ISSN : 0045-3102
- Online ISSN : 1468-263X
- Title id : bjsw
- Title URL : <http://bjsw.oxfordjournals.org>

Toutes les revues de OUP suivent cette architecture : le nom abrégé (Title id) permet de créer un sous-domaine de < oxfordjournals.org >. Ainsi, tous les contenus liés au British journal of social work ont pour racine URL < <http://bjsw.oxfordjournals.org> >.

La page d'accueil est claire et concise, et propose une navigation immédiate vers le numéro en cours, les archives ou numéros précédents, un outil de recherche, le comité de rédaction et les contacts actualisés. Elle intègre des liens vers les pages d'informations essentielles du journal (objectifs et champs d'application de la revue, comité de rédaction, instructions aux auteurs, système de contrôle qualité, engagements sur le libre accès, éthique éditoriale et plagiat, etc.).

La revue est disponible article par article avec un URL unique par article et un seul fichier PDF par document. Le chemin vers les articles permet de comprendre la granularité des documents : les visiteurs humains comme les robots d'indexation peuvent identifier « naturellement » la structure du site et l'articulation revue > volume > numéro > article (notice) > article (abstract / texte intégral HTML / texte intégral PDF).

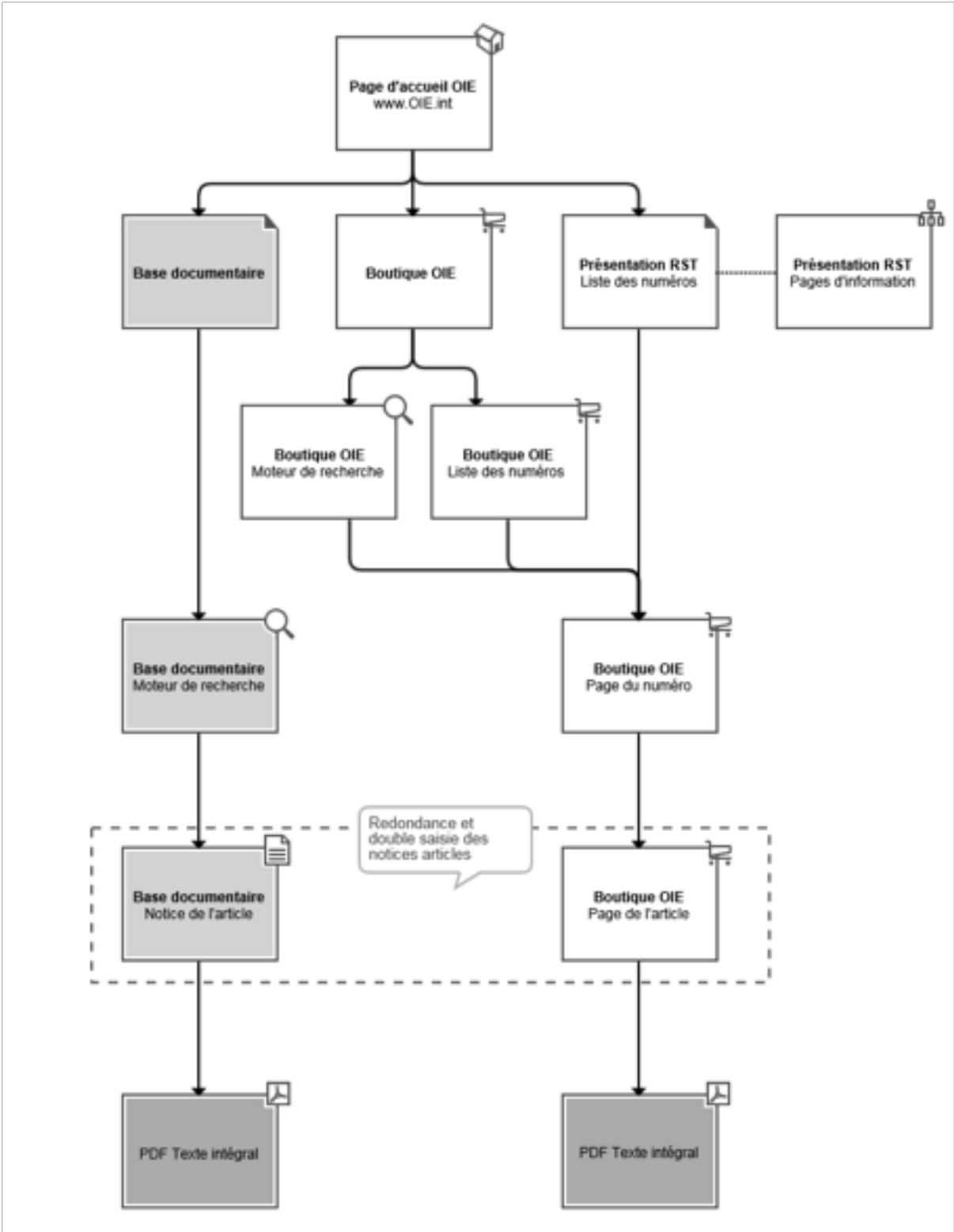
Cette architecture peut être utilement comparée à celle du site de l'OIE à l'été 2015. La liste ci-dessous présente la navigation permettant d'accéder au fichier PDF d'un article depuis la page d'accueil du site :

- Étape 1, page d'accueil du site OIE < <http://www.oie.int/fr/> >
- Étape 2, page Publications et documentation < <http://www.oie.int/fr/publications-et-documentation/information-generale/> >
- Étape 3, page *Revue scientifique et technique* < <http://www.oie.int/fr/publications-et-documentation/revue-scientifique-et-technique-acces-libre/liste-des-numeros/> >
- Étape 4, page du numéro, dans la partie marchande du site < [http://web.oie.int/boutique/index.php?page=ficprod&id\\_produit=1307&fichrech=1&lang=fr](http://web.oie.int/boutique/index.php?page=ficprod&id_produit=1307&fichrech=1&lang=fr) >
- Étape 5, page de l'article, dans la partie marchande du site < [http://web.oie.int/boutique/index.php?page=ficprod&id\\_prec=1307&id\\_produit=1368&lang=fr&fichrech=1](http://web.oie.int/boutique/index.php?page=ficprod&id_prec=1307&id_produit=1368&lang=fr&fichrech=1) >
- Étape 6, fichier PDF du texte intégral de l'article (fac-similé de l'édition imprimée) < <http://web.oie.int/boutique/extrait/04dallavilla3946.pdf> >

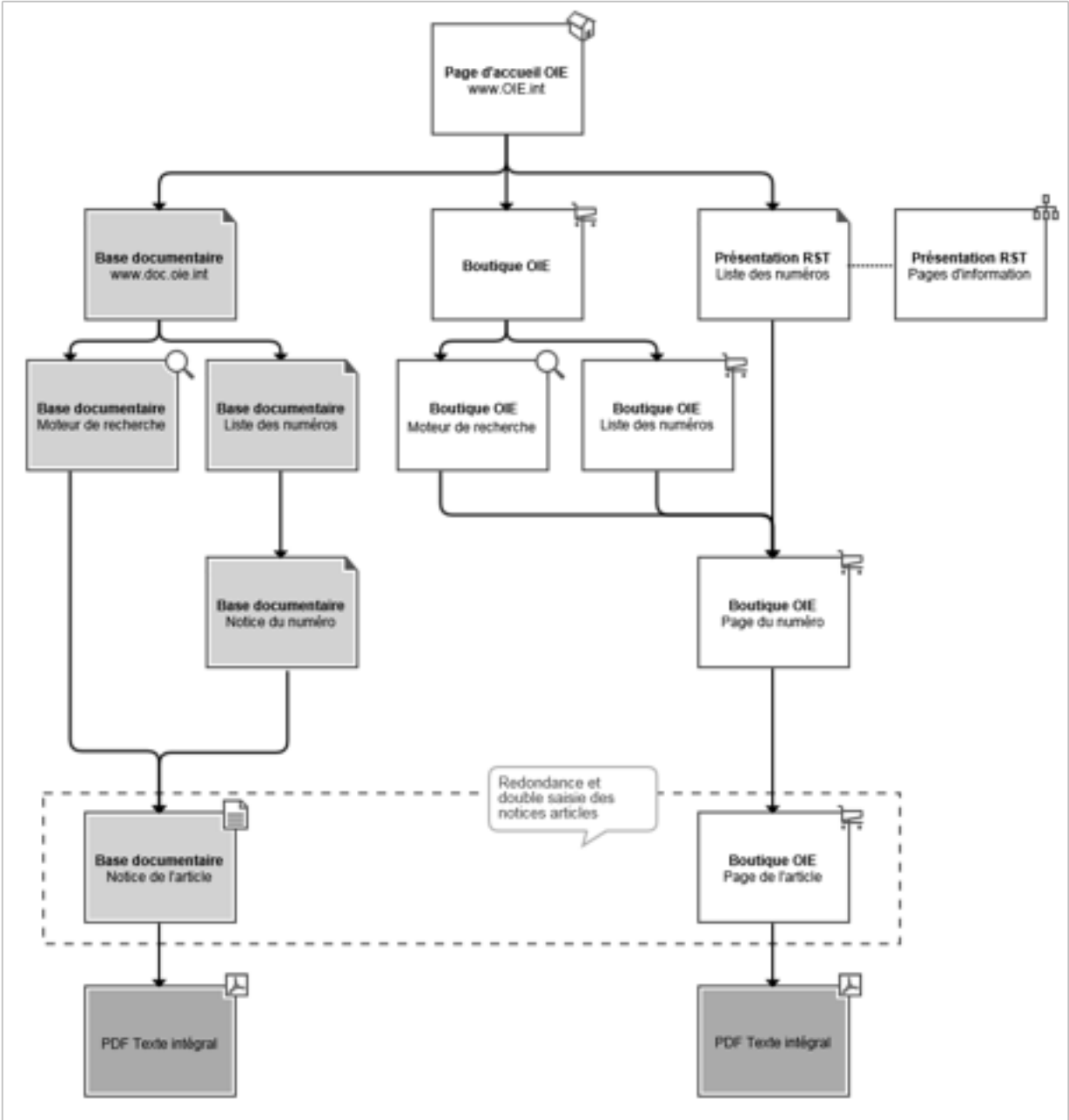
Bien qu'il ne s'agisse pas du seul chemin possible (voir Figure 3), et sans tenir compte de la question du multilinguisme, on constate que le nommage des URL ne correspond pas à une arborescence facilement intelligible par un humain ou une machine.

D'autre part, il existe une redondance entre les notices des articles dans la partie marchande du site et les notices dans la base documentaire.

**Figure 3 : Navigation pour accéder au texte intégral d'un article de la *Revue scientifique et technique* (état au 31 août 2015)**



**Figure 4 : Navigation pour accéder au texte intégral d'un article de la *Revue scientifique et technique* après mise en ligne du nouveau portail documentaire de l'OIE**



Le nouveau portail documentaire proposera une navigation structurée selon le niveau de granularité des documents. La redondance entre la partie marchande du site et le portail documentaire ne sera pas éliminée : chacune des deux sections continuera de gérer des notices distinctes, avec accès au texte intégral, pour un même article.

Comme première alternative, et sur le modèle de Oxford University Press, il pourrait donc être intéressant de créer un sous-domaine du site, dédié à la revue (par exemple < <https://rst.oie.int> >). Ce point d'entrée unique pourrait utiliser certains contenus du nouveau portail documentaire de l'OIE tels que le moteur de recherche ou les pages Volume > Numéro > Article (soit par un simple lien ou au travers d'une intégration plus poussée, ce qui nécessiterait du développement). En complément, il serait envisageable de supprimer les notices des articles sur la partie marchande du site (ne conserver que les notices de numéros), tout en renforçant les liens croisés entre cette partie et le portail documentaire. Cela permettrait de supprimer la redondance des notices et d'optimiser le référencement, avec une seule URL par document.

Cette recommandation visait à nourrir la réflexion sur l'organisation générale du site, ce qui nécessite une concertation assez large. Pour cette raison, j'ai considéré qu'il n'y avait pas lieu de reproduire ici la modélisation proposée – déjà obsolète au vu des réflexions en cours – mais d'en livrer les lignes générales. Plusieurs autres architectures favorisant le référencement naturel sont envisageables. En complément, la consultation du guide de bonnes pratiques *NISO RP-16-2013, PIE-J: Presentation & Identification of E-Journals* [16, NISO] permet de découvrir de nombreux exemples commentés de sites de revues académiques.

### **3.1.2 Instructions aux robots d'indexation**

Afin d'optimiser le parcours des robots à l'intérieur du site de l'OIE, et conformément aux instructions de Google Scholar pour le référencement (< <https://scholar.google.com/intl/en/scholar/inclusion.html#crawl> >, consulté le 15 septembre 2015), j'ai préconisé la modification du fichier robots.txt placé à la racine du site de l'OIE pour autoriser ou interdire l'indexation de certaines pages.

Il convient d'ajouter les instructions :

User-agent: \*

Disallow: /add\_cart.php

Cette modification interdira l'indexation des paniers d'achat dans la partie marchande, qui surcharge inutilement le parcours des robots et entraîne mécaniquement une dégradation du classement du site par Google.

### **3.1.3 Enrichissement sémantique des pages du portail documentaire**

Le nouveau portail documentaire permet d'enrichir sémantiquement les pages Web, afin de les rendre intelligibles par les moteurs de recherche ou les outils de citation.

Pour les pages des notices bibliographiques, Alexandria permet d'enrichir les balises *meta* du code HTML (*Hypertext Markup Language*, format de données conçu pour représenter les pages Web) avec des métadonnées. Je n'ai pas retenu le vocabulaire proposé sur le site < [www.schema.org](http://www.schema.org) >, site lancé en 2011 à l'initiative commune de

Google, Yahoo, Bing et Yandex pour enrichir sémantiquement le contenu des pages Web avec des vocabulaires. Ces vocabulaires permettent notamment aux moteurs de recherche d'appréhender des informations précises. En effet, l'usage du vocabulaire PublicationVolume (< <https://schema.org/PublicationVolume> >, consulté le 15 septembre) ne m'a semblé ni suffisamment répandu ni pertinent au regard du besoin. Après avoir analysé les instructions aux éditeurs pour le référencement naturel des publications périodiques dans Google Scholar (< <https://scholar.google.com/intl/en/scholar/inclusion.html> >, consulté le 15 septembre 2015) et étudié le code source des pages internet de plusieurs revues académiques (dont la revue *Parasite* et le site < <http://www.openedition.org> >), j'ai choisi d'exprimer les métadonnées dans les formats Dublin Core, citation\_ et PRISM (< [http://www.prismstandard.org/guides/pamw\\_guide\\_v1.0.htm](http://www.prismstandard.org/guides/pamw_guide_v1.0.htm) >, consulté le 15 septembre 2015).

Concrètement, Alexandria permet d'associer aux étiquettes des champs de métadonnées leurs équivalences exprimées dans d'autres formats, appelés « schémas ». Une fois effectuée cette cartographie des équivalences (mapping), il devient possible de sélectionner un schéma pour que chacune de ses métadonnées génèrent une balise *meta* dans l'en-tête (*header*) du code HTML de la page. Ainsi, les robots d'indexation seront capables de « comprendre » la page sur laquelle ils se trouvent : les métadonnées, exprimées selon des normes reconnues, sont interprétables par une machine. C'est le cas par exemple lors des opérations de scrapping (extraction automatique des métadonnées) par des outils de citation comme Zotero (Figure 5). Les utilisateurs de ces outils peuvent alors importer et partager facilement les références bibliographiques des documents consultés sur le site de l'OIE.

**Figure 5 : Métadonnées bibliographiques identifiées par l'outil de citation Zotero à partir des balises *meta* de pages HTML. Comparaison des données identifiées sur la notice d'un article dans la partie marchande du site (à gauche, page sans balise meta) et sur la notice équivalente dans le portail documentaire (à droite, après ajout des schémas Dublin Core, citation\_ et PRISM dans les balises *meta*)**

The image displays two side-by-side screenshots of the Zotero citation tool interface, comparing metadata extraction from a webpage (left) and a journal article (right).

**Left Window (Page Web):**

- Info** | Notes | Marqueurs | Connexe
- Type de document:** Page Web
- Titre:** OIE - Organisation Mondiale de la santé animale
- Auteur:** (Nom), (Prénom)
- Résumé:**
- Titre du site Web:**
- Type de site Web:**
- Date:**
- Titre abrégé:**
- URL:** <http://web.oie.int/boutique/index.php?page...>
- Consulté le:** 13/8/2015 10:33:52
- Langue:**
- Autorisations:**
- Extra:**
- Date d'ajout:** 13/8/2015 10:33:52
- Modifié le:** 13/8/2015 10:33:52

**Right Window (Article de revue):**

- Info** | Notes | Marqueurs | Connexe
- Type de document:** Article de revue
- Titre:** Ticks as vectors: taxonomy, biology and ecology
- Auteur:** A. Estrada-Peña
- (...) Résumé:** Ticks are prominent parasites and competent vectors of pathogens that affect both humans and animals. This review outlines and illustrates the main features of the morphology of ticks of the families Ixodidae and Argasidae, and summarises the basic components of their life cycles. It focuses mainly on development processes and mortality
- Publication:** Rev. sci. tech. Off. int. Epiz.
- Volume:** 34
- Numéro:** 1
- Pages:** pp. 53-65
- Date:** 04/2015
- Collection:**
- Titre de la coll.:**
- Texte de la coll.:**
- Abrév. de revue:**
- Langue:** ANGLAIS
- DOI:**
- ISSN:** 0253-1933
- Titre abrégé:** Ticks as vectors
- URL:** <http://217.19.49.201:8080/dyn/portal/in...>
- Consulté le:** 14/8/2015 16:41:26
- Archive:**
- Loc. dans l'archive:**
- Catalogue de bibl.:** 217.19.49.201:8080
- Cote:**
- Autorisations:** © OIE (World Organisation for Animal Health), 2015
- Extra:**
- Date d'ajout:** 14/8/2015 16:41:26



**Figure 6 : Balises *meta* d'une notice sur le portail documentaire de l'OIE (afin de garantir une bonne lisibilité, il s'agit d'une copie écran sans retour à la ligne automatique, ce qui permet de voir l'ensemble des balises malgré le texte important contenu par exemple dans les résumés)**

```

<head>
<meta http-equiv="Content-Type" content="text/html; charset=UTF-8" />
<meta http-equiv="X-UA-Compatible" content="IE=EmulateIE8" />
<meta name="generator" content="Alexandrie 7" />
<meta name="DC.title" content="Ticks as vectors: taxonomy, biology and ecology" />
<meta name="DC.creator" content="ESTRADA-PE&Ntilde;A A." />
<meta name="DC.subject" content="MALADIE A TRANSMISSION VECTORIELLE;TIQUE;BIOLOGIE;CLIMAT;HOTE;ECOLOGIE;MALADIE TRANSMISE PA
<meta name="DC.description" content="Ticks are prominent parasites and competent vectors of pathogens that affect both human
<meta name="DC.contributor" content="OIE" />
<meta name="DC.date.issued" content="04/2015" />
<meta name="DC.identifier" content="0253-1933" />
<meta name="DC.source" content="Rev. sci. tech. Off. int. Epiz." />
<meta name="DC.source" content="OIE" />
<meta name="DC.language" content="ANGLAIS" />
<meta name="DC.relation.isPartOf" content="New developments in major vector-borne diseases. Part One: An overview" />
<meta name="DC.relation.isPartOf" content="Rev. sci. tech. Off. int. Epiz." />
<meta name="DC.relation.isPartOf" content="34" />
<meta name="DC.relation.isPartOf" content="1" />
<meta name="DC.rights" content="&copy; OIE (World Organisation for Animal Health), 2015" />
<meta name="DC.description.abstract" content="Ticks are prominent parasites and competent vectors of pathogens that affect b
<meta name="DC.description.abstract" content="Las garrapatas son importantes par&acute;sites y eficaces vectores de pat&oc
<meta name="DC.description.abstract" content="Les tiques sont des parasites importants de l&#8217;homme et des animaux et de
<meta name="prism.alternateTitle" content="Les tiques en tant que vecteurs : taxonomie, biologie et &acute;ecologie" />
<meta name="prism.alternateTitle" content="Las garrapatas como vectores: taxonom&iacute;a, biolog&iacute;a y ecolog&iacute;a
<meta name="prism.keyword" content="MALADIE A TRANSMISSION VECTORIELLE;TIQUE;BIOLOGIE;CLIMAT;HOTE;ECOLOGIE;MALADIE TRANSMISE
<meta name="prism.originPlatform" content="Document OIE" />
<meta name="prism.publicationDate" content="04/2015" />
<meta name="prism.publicationName" content="Rev. sci. tech. Off. int. Epiz." />
<meta name="prism.copyright" content="&copy; OIE (World Organisation for Animal Health), 2015" />
<meta name="prism.issueName" content="New developments in major vector-borne diseases. Part One: An overview" />
<meta name="citation_title" content="Ticks as vectors: taxonomy, biology and ecology" />
<meta name="citation_author" content="ESTRADA-PE&Ntilde;A A." />
<meta name="citation_publication_date" content="04/2015" />
<meta name="citation_journal_title" content="Rev. sci. tech. Off. int. Epiz." />
<meta name="citation_issn" content="0253-1933" />
<meta name="citation_volume" content="34" />
<meta name="citation_issue" content="1" />
<meta name="citation_firstpage" content="pp. 53-65" />
<meta name="citation_abstract" content="Ticks are prominent parasites and competent vectors of pathogens that affect both hu
<meta name="citation_abstract" content="Las garrapatas son importantes par&acute;sites y eficaces vectores de pat&ocacute;ge
<meta name="citation_abstract" content="Les tiques sont des parasites importants de l&#8217;homme et des animaux et des vect

```

## 3.2 Identifiants pérennes

L'attribution de DOI (*Digital Object Identifier*) aux publications de l'OIE – notamment aux articles de la *Revue scientifique et technique* –, garantirait :

- la pérennité et la facilité d'accès aux articles, même en cas de migration des documents de l'OIE sur internet (comme cela va être le cas lors du lancement du nouveau portail documentaire : les URL actuels des fichiers PDF deviendront inactifs) ;

- la conformité vis-à-vis des standards et bonnes pratiques en matière d'interopérabilité des métadonnées pour les publications académiques ;
- une intégration et une visibilité plus forte dans l'écosystème des publications STM et, au final, auprès du public naturel des publications (les DOI aident à générer du trafic vers le contenu en le rendant détectable et plus facile à citer) ;

Une étude était nécessaire pour comprendre et préparer la mise en œuvre de DOI à l'OIE.

### **3.2.1 DOI**

Un DOI est un identifiant qui est l'émanation d'un autre identifiant : l'URI. Sur internet, les URI (*Uniform Resource Identifiers*) permettent d'identifier une ressource. On distingue généralement deux formes d'URI : les URL (*Uniform Resource Locators*) et les URN (*Uniform Resource Names*). Un URL – qui correspond à un lien hypertexte – identifie l'emplacement d'une ressource, plutôt que la ressource elle-même. Lorsqu'une ressource est déplacée, par exemple mise sur un autre serveur internet, son URL devient obsolète (on parle souvent de lien mort). À l'inverse, les URN identifient les ressources elles-mêmes, indépendamment de leur emplacement.

Un DOI (*Digital Object Identifier*) est un URI qui identifie à la fois l'emplacement d'une ressource et la ressource elle-même. Il s'agit d'un identifiant unique et pérenne, au centre d'un système d'identification de ressources numériques. Ce système est lui-même fondé sur le système Handle < <https://www.handle.net> > (antérieur au protocole TCP/IP), dont il emprunte un sous-domaine matérialisé par le préfixe « 10 » avec lequel débute tout DOI. Un protocole de résolution de lien assure une redirection

automatique en cas de changement d'adresse du document sur internet : cela évite les liens morts pour les articles cités dans un article. Cette question de l'obsolescence des liens (*reference rot*) est particulièrement documentée ces dernières années et toucherait un article sur cinq en STM, ce qui signifie qu'il est impossible de retrouver le contexte internet qui entourait les publications concernées [11, KLEIN].

Il est important de noter que les DOI ne sont pas les seuls types d'identifiants pérennes. D'autres systèmes existent tels que ARK, PURL, Handle [3, CLAERR, pp. 363-367, et 1, BERMES, pp. 76-77]. Il convient également de se garder de toute pensée magique sur le « solutionnisme » des DOI : plutôt que des liens pérennes, il faudrait parler de liens pérennisables. Sans volonté forte de maintenir les métadonnées à jour, il est probable que les liens deviendront tôt ou tard inactifs. Cela implique le maintien du système lui-même (frappé par une panne géante en janvier 2015, liée au non renouvellement du nom de domaine doi.org [10, CROSSTECH]) et la sensibilisation des éditeurs pour la mise à jour en temps réel des métadonnées.

Un autre péril concerne la modification intempestive du contenu des documents (*content drift*) qui a conduit certains acteurs tels que Wikipedia à se détourner des DOI au profit de l'archivage de captures de pages Web.

Magré cela, les DOI sont le système à privilégier par les éditeurs pour identifier les objets (articles, livres numériques, jeux de données, fichiers sonores, vidéo, etc.) sur le Web.

Le DOI se compose d'un préfixe et d'un suffixe séparés par un « / ». Le préfixe commence toujours par « 10 » suivi de quatre chiffres ou plus. Le préfixe est attribué par une agence DOI agréée par la DOI Fondation. Il identifie l'éditeur (mais il est possible à un éditeur de demander plusieurs préfixes, par exemple un par revue). Le suffixe est à l'inverse attribué par l'éditeur lui-même. C'est une chaîne de caractères alphanumériques insensibles à la casse, qui ne peut inclure que les caractères approuvés : "az", "AZ", "0-9" et "-.\_; () /". Il est recommandé de ne pas choisir de suffixe trop long.

La norme ISSN (International Standard Serial Numbering, pour l'identification unique des titres périodes, ISO 3297), recommande un suffixe DOI incluant l'ISSN (avec trait d'union), précédé des lettres minuscules « ISSN », par exemple pour la *Revue scientifique et technique* : 10.5930/issn.0253-1933

Selon les recommandations de l'agence Crossref (< <http://www.crossref.org/> >), il est préférable d'exprimer un DOI sous forme de lien hypertexte. Un lien DOI se compose de deux parties : le répertoire URL DOI < <http://dx.doi.org/> >, et le DOI lui-même. Lorsqu'il est combiné ainsi, le DOI est de fait un lien hypertexte. Toujours dans le cas de la *Revue scientifique et technique*, nous aurons : < <http://dx.doi.org/10.5930/issn.0253-1933> >.

L'exemple ci-dessus correspond à un DOI attribué au niveau du titre du périodique. Il est en effet recommandé d'attribuer un identifiant par niveau de granularité (par exemple : la revue > le volume > le numéro > l'article).

Dès que les DOI (et leurs métadonnées) sont déclarés auprès d'une agence d'enregistrement, ils peuvent être repris et cités. L'éditeur doit donc immédiatement créer des pages de réponse actives pour recevoir les liens. Habituellement, une page de réponse (*landing page*) correspond à la notice bibliographique du document avec son résumé, librement accessibles. De là, des liens – idéalement interprétables par les machines comme par les humains – proposent l'accès au texte intégral (avec ou sans authentification, en version HTML ou PDF, etc.). Si le texte complet est disponible sans frais, les utilisateurs peuvent y accéder immédiatement.

L'affectation de DOI distincts pour l'article et ses traductions est recommandée par CrossRef mais facultative. Après avoir étudié les spécificités de la *Revue scientifique et technique*, je recommande plutôt une seule attribution de DOI, avec une mention des traductions dans les métadonnées et une page de réponse unique, qui renverra vers l'ensemble des versions de l'article.

D'après les chiffres de la International DOI Foundation (IDF) datés du 22 juin 2015 (< <http://www.doi.org/factsheets/DOIKeyFacts.html> >, consulté le 30 septembre 2015), il existe plus de 114 millions de DOI attribués (dont 75 millions par la seule agence CrossRef), avec un taux de croissance annuel de 19%. Ceux-ci sont liés à plus de 15 000 préfixes DOI (taux de croissance annuel : 27%). Le système dans son ensemble permet la résolution de plus de 1,2 milliards de liens par an.

### **3.2.1.1 Comparaison des agences CrossRef et DataCite**

Deux agences DOI étaient envisagées par l'OIE : CrossRef et DataCite. Il convenait de comparer chaque agence par rapport aux besoins de l'Organisation.

CrossRef et DataCite sont membres de la International DOI Foundation (IDF), instance dirigeante du système DOI. Il s'agit de services payants, fondés sur des technologies propriétaires. Ces agences offrent aux éditeurs y souscrivant les avantages suivants :

- les objets (publications ou jeux de données) sont immédiatement « citables » dès l'attribution du DOI.
- ces systèmes garantissent un accès cohérent aux objets grâce à des liens stables. CrossRef est une agence dédiée aux publications académiques. DataCite, représentée en France par l'INIST, est davantage dédiée aux jeux de données.

**Tableau 3 : Comparaison des agences DOI DataCite et CrossRef**

	<b>DataCite / INIST</b>	<b>CrossRef</b>
<b>Organisation</b>	Consortium international dont l'administrateur général est la Bibliothèque Nationale de science et technologie allemande (TIB). L'Inist-CNRS est l'institution membre de DataCite pour la France.	Consortium international. Initiative de l'association d'éditeurs scientifiques PILA (Publishers International Linking Association) destinée à améliorer la visibilité, le signalement, la découverte et la citation des articles de périodiques en ligne par des services de « références croisées ».
<b>Positionnement</b>	<b>Jeux de données</b> (objets issus de la recherche)	<b>Publications</b> (faciliter l'accès aux articles et aux résultats de la recherche mondiale)
<b>Références</b>	<b>Centres de recherche.</b> Exemples : CERN, Microsoft Research, INIST... Chiffres non communiqués.	<b>Editeurs et services de publication.</b> 4500 organisations, quasi-intégralité des éditeurs STM (géants mondiaux – Elsevier, Wiley, OCDE – comme éditeurs plus modestes – EDP Sciences –).  74 million d'identifiants déclarés (données juin 2015)
<b>Coût</b>	Droit annuel : 180,00€ Coût variable par DOI : 0,00€  Total année 1 : 180,00€ Total années suivantes : 180,00€	Droit annuel : 275,00\$ Coût variable par DOI, à payer une seule fois lors de la déclaration : Ressource publiée depuis 2013 : 1,00\$ Ressource publiée avant 2013) : 0,15\$  <b>Estimation du coût des déclarations DOI pour la seule <i>Revue scientifique et technique</i></b> Total année 1 : 787,16\$ (en supposant que tous les anciens articles + ceux de

	DataCite / INIST	CrossRef
		l'année 1 soient référencés) Total années suivantes : 365,00\$ (coût abonnement + référencement de tous les articles des 3 numéros de l'année)
<b>OAI-PMH</b>	<b>Oui.</b> Peu utilisé : la plupart des références concerne des objets non textuels.	<b>Oui.</b> Moissonnage du serveur OAI (payant) par Pubmed, CAB Abstracts, Zotero, résolveurs de lien, consortia et bases de connaissance (Serials Solutions 360, EBSCO AtoZ, ExLibris SFX, LinkSolver/Links@Ovid, EBSCO Discovery Service, Proquest Summon, ExLibris Primo Central)
<b>Citations croisées</b>	<b>Non.</b> Aucune obligation pour les organisations affiliées de citer les DOI pour désigner un jeu de données.	<b>Oui (impérative)</b> Les éditeurs affiliés ont l'obligation d'intégrer les DOI dans les citations. Une API de recherche est disponible à cet effet.  <b>Fonctionnalité « Cited By »</b> Pour enrichir les notices de façon dynamique ("Cet article a été cité par"), les éditeurs ont la possibilité de demander au système d'extraire la liste des articles citant l'une de leurs publications.
<b>Devise facturation</b>	Euro	Dollar
<b>Langue service client et support</b>	Anglais et français	Anglais

	DataCite / INIST	CrossRef
<b>Obligations OIE</b>	<p><b>Métadonnées à fournir</b></p> <p>XML impératif (selon schéma .xsd DataCite)</p> <p><i>Tolérance : non</i></p>	<p><b>Métadonnées à fournir</b></p> <p>XML impératif (selon schéma .xsd CrossRef)</p> <p><i>Tolérance : non</i></p> <p><b>Citations avec DOI</b></p> <p>Pour chaque référence citée dans un article de la <i>Revue scientifique et technique</i>, extraire de la base de données CrossRef le DOI menant vers l'article cité (s'il existe) et le faire figurer dans la bibliographie.</p> <p><i>Tolérance : 18 mois à partir de l'ouverture du compte</i></p> <p><b>COUNTER compliant</b></p> <p>Le projet COUNTER facilite l'enregistrement, l'échange et l'interprétation des données d'utilisation électronique. Des standards et des protocoles ouverts et internationaux permettant aux fournisseurs de contenu de produire des statistiques d'utilisation cohérentes, crédibles et compatibles.</p> <p><i>Tolérance : oui si peu de publications</i></p>
<b>Délai</b>	Ouverture du compte et attribution du préfixe DOI à réception du paiement du droit annuel.	Ouverture du compte et attribution du préfixe DOI à réception du paiement du droit annuel. Puis facture trimestrielle selon dépôts de DOI.



Après étude des spécificités de chaque agence, j'ai recommandé CrossRef, agence plébiscitée par les éditeurs internationaux et adaptée à la finalité du besoin exprimé :

- pérennité des URI liées aux publications de l'OIE ;
- granularité variable et recomposable des objets identifiés (collection, livre, revue, chapitre, article, tableau, liste de référence...);
- visibilité accrue (métadonnées CrossRef réutilisées par PubMed, consortia de bibliothèques universitaires, bases de connaissance, résolveurs de lien, Zotero) ;
- maîtrise des éléments de citation dans les publications (en choisissant l'URI de destination, l'OIE maîtrise les citations : le lien renvoie vers la page souhaitée par l'OIE) ;
- performance accrue des statistiques bibliométriques (les articles sont correctement identifiés) ;
- intégration OpenURL. OpenURL nécessite une source d'identifiants et de métadonnées. Plutôt qu'organiser le site de l'OIE pour qu'il soit conforme à la norme OpenURL, ce qui impliquerait des développements assez conséquents, CrossRef fournit un point unique d'accès sur le réseau pour la création et la redirection ultérieure des liens OpenURL ;
- suivi détaillé des citations dans les publications savantes (fonction « Cited By », fonction DET DOI Event Tracker) et autres services associés (CrossCheck pour lutter contre le plagiat et CrossFund sur le financement de la recherche).

Les deux premiers points sont possibles avec DataCite, les points suivants demandent de choisir CrossRef. Le coût de CrossRef est moins avantageux.

### **3.2.1.2 Travailler avec CrossRef**

Après avoir recommandé l'agence DOI CrossRef, il m'a semblé utile d'approfondir les points d'attention à surveiller lors du lancement de la démarche.

#### *3.2.1.2.1 Calendrier prévisionnel*

La durée de déploiement estimée est de un mois à compter du choix de l'agence (15 jours pour la phase 2 et 15 jours pour les suivantes) :

1/ Choix de l'agence DOI

2/ Ouverture du compte et attribution du préfixe DOI

3/ Attribution des DOI (traitement par lots dans Alexandrie)

4/ Extraction du fichier XML (DOI, URL et métadonnées liées) :

- Paramétrage de la transformation XML pour conformité avec schéma cible [1 à 2 jours]
- Traitement par lot des attributions de DOI dans Alexandrie (requête SQL) [0,5 jour]

5/ Envoi du fichier XML à l'agence

### 3.2.1.2.2 Principaux services impactés

La pérennité est une décision et un investissement qui engage de nombreuses parties prenantes, avec lesquelles il est nécessaire de se concerter. Dans la configuration actuelle des rôles et responsabilités, la simple attribution de DOI impactera notamment les services suivants :

Systemes d'information :

- Garantir la conformité des fichiers de dépôt ;
- Maintenir un accès FTP pour le transfert des fichiers XML.

Documentation :

- Sélectionner les ressources qui feront l'objet d'une attribution de DOI ;
- Garantir l'envoi régulier des fichiers de dépôt ;
- Définir – en concertation avec les autres parties prenantes – la structure des suffixes DOI ;
- Veiller au renouvellement des droits annuels.

Publications :

- Attribuer un DOI à chaque article de la *Revue scientifique et technique*, au plus tôt du processus éditorial (mais veiller à ce que le DOI n'identifie que la version publiée) ;

- Récupérer et citer les DOI de toutes les références dans les bibliographies des articles de la *Revue scientifique et technique* (voir Annexe 1).

### 3.2.1.2.3 Dépôts XML

Pour chaque ressource enregistrée auprès de CrossRef, l'éditeur crée un DOI unique, en ajoutant un suffixe de son choix au préfixe attribué par CrossRef, et lui associe un URL.

CrossRef enregistre le DOI et son URL dans un répertoire DOI central (développé et maintenu par la International DOI Foundation). Le système fonctionne en interrogeant le répertoire pour connaître l'URL en cours de validité pour le DOI. Quand un éditeur modifie l'emplacement d'une ressource, le nouvel URL n'est à signaler qu'à CrossRef, qui met à jour le répertoire central.

Pour faire enregistrer une ressource, l'éditeur transmet un fichier XML horodaté qui associe des métadonnées de citation et le DOI. Le fichier doit être conforme à un schéma spécifique, exprimé au format .xds < <http://doi.crossref.org/schemas/crossref4.3.6.xsd> > (les versions supérieures ou égales à 4.3.0 sont actuellement acceptées pour les dépôts < [http://www.crossref.org/help/schema\\_doc/4.3.6/4.3.6.html](http://www.crossref.org/help/schema_doc/4.3.6/4.3.6.html) >, consulté le 30 septembre 2015).

J'ai préparé un fichier de transformation XSLT pour permettre l'exportation de métadonnées depuis Alexandria vers un fichier XML conforme. Je reviendrai sur le schéma CrossRef dans la partie consacrée au XML.

### 3.2.2 PMID (PubMed ID)

Le PMID (PubMed ID) est un identifiant unique attribué à un article après parution et validation des métadonnées par PubMed. Cette condition et ce décalage ne permettent pas de substituer les PMID aux DOI. D'autre part, sur PubMed, de nombreuses ressources numériques identifiées par un PMID ne sont pas associées à un lien hypertexte (c'est le cas actuellement de la *Revue scientifique et technique*). Pour autant, les PMID sont utilisés par les principaux outils de citation et résolveurs de lien pour identifier une ressource en ligne, en complément des DOI (l'accès aux métadonnées est gratuit sur PubMed, payant sur CrossRef).

J'ai donc recommandé d'intégrer les PMID aux notices bibliographiques du portail documentaire de l'OIE. Pour cela, j'ai construit une table de correspondance entre les PMID des 1832 articles de la revue référencés dans PubMed et leurs identifiants dans Alexandrie. Il s'agissait de confirmer par exemple que la notice < <http://www.ncbi.nlm.nih.gov/pubmed/?term=25812225> > (consultée le 15 septembre 2015) correspondait bien à la référence 32126 dans Alexandrie. Des choix de référencements dans Alexandrie (années 1994-1996 et 1998) et quelques incertitudes liées aux informations présentes dans PubMed expliquent l'essentiel des correspondances non trouvées.

À partir de cette table, un traitement par lots (fondé sur une requête SQL) permettra l'enrichissement du portail documentaire (opération planifiée à une date postérieure à la mission de stage, octobre 2015).

D'autre part, pour obtenir une attribution de PMID plus rapide, j'ai recommandé de fournir à PubMed les métadonnées des nouveaux articles, au plus tôt de leur publication. Chaque article fera alors l'objet d'une notice minimale, avant indexation complète par MEDLINE. Cette première étape permettra d'intégrer le programme

LinkOut, qui autorise le transfert à PubMed des DOI et liens hypertextes liés à un PMID.

J'ai là aussi préparé un fichier de transformation (au format XSLT) pour permettre l'exportation de métadonnées, depuis Alexandria vers un fichier XML conforme à la DTD (*Document Type Definition*, fichier permettant d'écrire les définitions d'un document XML, solution antérieure aux schémas XML) de PubMed. J'y reviendrai dans la partie consacrée au XML.

Pour LinkOut, il faudra fournir, par transfert FTP, un fichier .csv avec une colonne pour les PMID et une colonne pour les URL ou DOI.

### 3.3 Exports XML

Alexandrie permet de créer des « schémas » pour associer aux étiquettes des champs de métadonnées leurs équivalences exprimées dans d'autres formats. Une fois effectuée cette cartographie des équivalences (*mapping*), dont le tableau 4 montre un exemple, il devient possible :

- de sélectionner un schéma pour que les métadonnées associées génèrent des balises *meta* dans l'en-tête (*header*) du code HTML des notices (voir infra 3.1.3. Enrichissement sémantique des pages du portail documentaire)
- d'exporter les métadonnées au format XML, dans le schéma souhaité.

**Tableau 4 : Equivalence en Dublin Core des étiquettes de champs Alexandrie**

<b>Etiquette de champs dans Alexandrie</b>	<b>Equivalence Dublin Core</b>
TITO	dc:title
TITRen	dcterms:alternative lang="en"
AUT	dc:creator
KEY	dc:subject
NOTILL	dc:description
EDIT	dc:publisher
INST	dc:contributor
DPARU	dc:date / dcterms:issued
LANG	dc:language
KGEO	dcterms:spatial
COPY	dc:rights
RESUen	dc:description / dcterms:abstract lang="en"
TITRfr	dcterms:alternative lang="fr"
TITRes	dcterms:alternative lang="es"
RESUfr	dcterms:abstract lang="fr"
RESUes	dcterms:abstract lang="es"
PAGE	dc:description
URL	dc:identifier
DOI	dc:identifier
PMID	dc:identifier
ISBN	dc:identifier
CCD	dcterms:isPartOf
PERIO	dcterms:isPartOf
VOL	dcterms:isPartOf
NPERIO	dcterms:isPartOf

J'ai exploité cette deuxième possibilité pour répondre aux besoins de fichiers XML conformes pour :

- alimenter l'entrepôt OAI-PMH ;
- répondre aux demandes de PubMed, Agricola et CrossRef ;
- générer des fichiers .txt afin de travailler sur les données d'Alexandrie en contournant la complexité des requêtes SQL (l'organisation des tables de données dans Alexandrie est peu intuitive).

Pour générer les balises *meta* de pages HTML, la structure d'Alexandrie commande de créer un schéma par format de métadonnées (Dublin Core, citation\_ et PRISM). Pour générer les fichier XML, je suis parti d'un schéma pivot, en Dublin Core, sur lequel j'ai appliqué des fichiers de transformation XSLT.

Concrètement, il s'est agi de créer des requêtes spécifiques selon l'étendue des données à exporter (par exemple, tous les articles issus de publications de l'OIE ou seulement ceux de la *Revue scientifique et technique*, sur une période donnée, etc.). À partir de ces requêtes, j'ai défini des stratégies d'export (sélection du répertoire où le fichier sera déposé, choix du format – .txt ou .xml –, fichier de transformation XSLT à appliquer, etc.). Enfin, chaque stratégie d'export est liée à une valeur cron qui détermine la fréquence et l'heure de l'export (par exemple « 0 15 12 ? \* MON#1 » pour un export mensuel, chaque premier lundi du mois, à 12h15).

Je souhaitais convertir les fichiers exportés d'Alexandrie en fichiers conformes aux instructions édictées pour l'OAI-PMH, CrossRef, PubMed (ces instructions sont elles-mêmes exprimées dans des fichiers XML, dans les langages DTD ou Schema). J'ai donc écrit plusieurs fichiers de transformation XSLT (*eXtensible Stylesheet Language*



*Transformations*). Ce langage de transformation XML permet de transformer le contenu et la forme d'un fichier XML.

Ainsi, à partir d'un schéma pivot dans Alexandria, je peux définir une requête spécifique sur les données à exporter et appliquer un fichier de transformation XSLT au document XML généré par Alexandria.

Pour écrire les trois fichiers XSLT (un premier pour PubMed, Agricola et Agris, qui utilisent une DTD identique, un autre pour CrossRef et un dernier pour l'OAI-PMH), j'ai dû dans un premier temps m'autoformer à ce langage et à XPath. Des bases moyennes en XML suffisent pour atteindre en trois ou quatre jours le niveau de compétence permettant de répondre au besoin. En effet, de nombreux exemples de codes sont disponibles sur le Web pour traiter les points les plus délicats. Il a également fallu choisir un éditeur de textes pour écrire le code. J'avais dans un premier temps envisagé un logiciel commercial dédié au XML et au XSLT (tel que Oxygen, qui propose une version d'évaluation gratuite pendant 30 jours) mais ces outils – certes très puissants – impliquaient un délai assez long pour la « prise en main » de l'interface. Au final, j'ai utilisé les logiciels Notepad++ (sous Windows) et jEdit qui dispose d'un plugin XML très performant (sous Mac OS).

Plutôt que lire pas-à-pas les DTD et schémas, je suis parti de fichiers XML valides. Balise après balise, j'ai transformé le fichier XML brut exporté d'Alexandrie afin de le rendre valide.

Les points délicats rencontrés dans l'écriture des fichiers XSLT étaient relativement similaires. En voici quelques exemples :

- le multilinguisme :

- création d'attributs linguistiques pour les balises des titres et des résumés

- modification des données linguistiques pour les rendre conformes à la norme ISO 639 2b

- les formats de date :

- modification du format de date pour le rendre conforme à la norme ISO 8601 (AAAA-MM-JJ)

- la concaténation de plusieurs champs :

- la balise <dcterms:isPartOf> regroupe plusieurs données : nom du numero, nom de la revue, volume et numero

- les noms

- séparation du prénom et du nom (une seule donnée dans Alexandrie)

- la pagination

- les valeurs pour les première et dernière pages d'un article sont demandées par PubMed et CrossRef (mais exprimées sous la forme « pp. XX-XX » dans Alexandrie)

On voit ici tout l'intérêt du langage XSLT pour la gestion des métadonnées. Il est possible de transformer les données pour répondre des besoins spécifiques et réguliers, sans modifier celles-ci dans la base documentaire. En effet, la qualité des données présentes dans Alexandrie tient en premier lieu à la stabilité des formats de saisie. Il était donc possible de s'appuyer sur cette stabilité (après en avoir éprouvé la solidité et la cohérence par un traitement par lots afin de traiter les anomalies éventuelles) pour générer de nouvelles données « à la volée ».

Pour faciliter la modification future des fichiers XSLT créés, par exemple pour tenir compte d'un changement dans Alexandrie ou dans les DTD ou schémas, j'ai écrit dans le corps des fichiers de nombreux commentaires qui explicitent la finalité des lignes de codes.

### **3.3.1 Remarques sur la DTD PubMed**

Le fichier XSLT que j'ai écrit permet de générer des fichiers XML valides par rapport à la DTD du service PubMed Data Provider.

Un outil de validation en ligne (< <http://www.ncbi.nlm.nih.gov/pubmed/citcheck> >, consulté le 15 septembre 2015) signale comme unique anomalie que l'ISSN de la revue est déjà référencé dans PubMed (ce n'est donc pas une anomalie pour la validité du fichier).

L'étape suivante sera de transmettre par FTP les données relatives au prochain numéro de la *Revue scientifique et technique* en informant PubMed de cet envoi par un mail parallèle, qui précisera :

- Filename: (nom du fichier)
- Title: (Titre de la revue)
- ISSN: (ISSN de la revue)
- Vol: (Volume)
- Iss: (Issue, numéro)
- URL: (URL de la revue)

En cas de validité, les données seront en ligne sur PubMed sous 48h, avant d'être ultérieurement complétées par les données de MEDLINE. Dès que le premier fichier aura été accepté et mis en ligne (et seulement à cette condition), l'OIE pourra participer au programme LinkOut et ainsi enrichir les notices PubMed avec des DOI ou des liens vers le site de l'OIE (fourniture de fichiers .csv listant les PMID et les DOI ou URL associés).

Dans une note interne sur les exports XML, j'ai transmis à l'OIE les mots de passe du compte FTP ouvert par PubMed ainsi que les contacts utiles.

**Figure 7 : Extrait du fichier XSLT pour les flux XML vers PubMed**

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <xsl:stylesheet version="1.0"
3 xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
4
5 <xsl:output method="xml" version="1.0" encoding="UTF-8" omit-xml-declaration="no" indent="yes"
6 doctype-public="-//NLM//DTD PubMed 2.6//EN"
7 doctype-system="http://www.ncbi.nlm.nih.gov/corehtml/query/static/PubMed.dtd"/>
8
9 <xsl:template match="/">
10
11
12 <!--Fichier XSLT créé par Romuald Verrier le 2015-08-20 d'après DTD XML PubMed Data Provider. Reserve exclusi
13
14
15
16
17
18 <ArticleSet><xsl:text>&#xA;</xsl:text>
19
20 <xsl:for-each select="./records/record">|
21
22 <Article><xsl:text>&#xA;</xsl:text>
23 <Journal><xsl:text>&#xA;</xsl:text>
24
25
26 <!--creation PublisherName ATTENTION invariable-->
27
28 <PublisherName>OIE World Organisation for Animal Health</PublisherName>
29 <xsl:text>&#xA;</xsl:text>
30
31
32 <!--creation JournalTitle ATTENTION invariable-->
33
34 <JournalTitle>Rev Sci Tech</JournalTitle>
35 <xsl:text>&#xA;</xsl:text>
36
37
38 <!--creation ISSN ATTENTION invariable-->
39
40 <Issn>0253-1933</Issn>
41 <xsl:text>&#xA;</xsl:text>
42
43 <!--creation Volume-->
44
45 <xsl:for-each select="./VOL">
46 <Volume>
47 <xsl:value-of select="."/>
48 </Volume>
49 <xsl:text>&#xA;</xsl:text>
50 </xsl:for-each>
51
52
53 <!--creation Issue-->
54
55 <xsl:for-each select="./NPERIO">
56 <Issue>
57 <xsl:value-of select="."/>
58 </Issue>
59 <xsl:text>&#xA;</xsl:text>

```

### **3.3.2 Remarques sur le schéma CrossRef**

Le fichier XSLT créé permet la conformité au schéma de dépôt CrossRef 4.3.4 (< [http://help.crossref.org/deposit\\_schema\\$citation](http://help.crossref.org/deposit_schema$citation) >, consulté le 15 septembre 2015).

Les dépôts XML consisteront à fournir à CrossRef les métadonnées des articles pour lesquels un DOI est déclaré.

En l'absence de contrat entre CrossRef et l'OIE, je n'ai pas pu faire valider les fichiers par CrossRef pendant le stage. Il restera par ailleurs à définir une source de données pour l'horodatage des fichiers déposés (balises <doi\_batch\_id> et <timestamp>).

## **3.4 Entrepôt OAI-PMH**

### **3.4.1 Le protocole OAI-PMH**

Le protocole OAI-PMH (*Open Archives Initiative – Protocol for Metadata Harvesting*) permet d'exposer des métadonnées sur internet, à destination d'une diversité d'acteurs. Il s'agit d'un cadre général pour la fédération de contenus sur internet, qui permet d'accroître la visibilité des ressources numériques, de reconstituer l'index de collections disponibles, d'enrichir des fonds numériques avec des ressources distantes, etc. Les métadonnées décrivent généralement des documents numériques et sont

exprimées dans le format Dublin Core (bien que le format MARC soit également très répandu).

Le protocole est libre d'utilisation et les spécifications de la version 2 (juin 2002) sont consultables sur le site de l'organisation initiatrice (< <http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm> >, consulté le 15 septembre 2015).

OAI-PMH est un dispositif asynchrone (les données sont mises à jour sur le serveur par des versements programmés à fréquence régulière et non en continu). Fondé sur les standards HTTP et XML, le protocole repose sur une communication de client (requête en HTTP) à serveur (réponse par un flux de données en XML). Il y a donc deux types d'acteurs :

- les fournisseurs de données, qui exposent leurs métadonnées sur un serveur internet (entrepôt OAI-PMH)
- les utilisateurs qui « moissonnent » (*harvest*) un ou plusieurs entrepôts OAI-PMH pour collecter les métadonnées selon leurs besoins

Un même acteur peut exposer ses métadonnées et moissonner celles des autres.

### **3.4.2 Le format Dublin Core**

Le format Dublin Core est un « schéma de métadonnées générique qui permet de décrire des ressources numériques ou physiques et d'établir des relations avec d'autres ressources (ISO 15836 depuis 2003) » [2, BOURDENET, p. 466].

Le Dublin Core non qualifié définit 15 éléments de base (*core elements*), facultatifs et répétables. Selon la présentation qu'en fait Emmanuelle Bermès [1, BERMES, p. 33] :

« le Dublin Core a été par la suite enrichi pour former une liste extensive de propriété et de classes, groupée sous l'appellation de Dublin Core Metadata Initiative (DCMI) Metadata Terms. Chaque propriété et chaque classe de cette liste de termes dispose d'une URI et d'informations sur sa logique et son usage, exprimés en RDFS. Dublin Core est dès lors compatible avec le Web sémantique. »

Le Dublin Core, qualifié ou non qualifié, est donc particulièrement adapté pour l'interopérabilité des métadonnées. Parce qu'il autorise une grande liberté dans le format des données, le DCMI est souvent considéré comme le « plus petit dénominateur commun » pour le partage d'informations [1, BERMES, pp. 32-37]. Mais cette faiblesse est avant tout la raison de son succès en tant que format minimal, au risque d'une faible qualité des données.

Pour cartographier les équivalences entre les étiquettes de champs dans Alexandrie et les éléments Dublin Core (voir tableau 4, dans infra 3.3. Exports XML), je me suis fondé sur un guide d'utilisation du Dublin Core à la BNF [13, BNF] ainsi que sur le site du DCMI (< <http://dublincore.org/documents/dcmi-terms/> >, consulté le 15 septembre 2015). En effet, l'étude comparative de quelques entrepôts OAI-PMH m'a permis de constater de grandes divergences d'interprétation dans l'utilisation des éléments DCMI, certains éditeurs leur rattachant des métadonnées sans rapport (par exemple le DOI du document décrit était parfois – à tort – associé à « dc:source », parfois à « dc:identifier »). En raison de ces contradictions, l'analyse des pratiques existantes ne pouvaient servir de guide et la consultation de sources fiables étaient donc indispensables.

Enfin, comme le Dublin Core peut être utilisé en HTML et en XML mais avec des syntaxes différentes. Par exemple, on écrira « DC.title » en HTML (au sein de la balise meta) et « dc:title » en XML. Cela m'a conduit à créer deux schémas spécifiques dans

Alexandrie : l'un pour générer les balises *meta* des pages du portail documentaire, l'autre pour les exports XML.

### **3.4.3 Installation d'un entrepôt OAI-PMH à l'OIE**

Pour l'OIE, la mise en place d'un entrepôt OAI-PMH permettait d'offrir un accès normalisé à l'ensemble des métadonnées des articles, pour leur référencement par une multitude d'acteurs (bibliothèques et centres de documentation, bases de connaissance, moteurs de recherche, etc.).

Alors que dans certains contextes le coût global d'un entrepôt OAI-PMH peut s'avérer prohibitif [12, SIREAU, pp. 69-80], la nouvelle version d'Alexandrie permet une mise en œuvre pour un coût marginal insignifiant. Une fois les données structurées en Dublin Core (utilisation des fonctions schémas et des requêtes d'export), il est possible – grâce à un connecteur intégré et gratuit – d'utiliser le logiciel libre jOAI (< [http://www.dlese.org/dds/services/joai\\_software.jsp](http://www.dlese.org/dds/services/joai_software.jsp) >, consulté le 15 septembre 2015) qui a la double fonction d'entrepôt et de moissonneur OAI-PMH. Conformément aux instructions de la société GB Concept (éditeur d'Alexandrie), j'ai installé jOAI sur le serveur d'application GlassFish. Le paramétrage est très intuitif.

J'ai créé deux sets (ou jeux de données), un pour le Bulletin, l'autre pour la *Revue scientifique et technique*. Dans un premier temps, les données étaient limitées aux notices postérieures à 2013. Il conviendra pour l'OIE d'élargir les requêtes Alexandrie pour présenter l'ensemble des notices et, éventuellement, créer un troisième set pour exposer ses autres publications. Il faudra aussi définir la fréquence d'actualisation optimale (un export hebdomadaire depuis Alexandrie semblait suffisant).



J'ai également proposé des textes trilingue pour la création d'une page dédiée à l'entrepôt OAI-PMH sur le prochain portail documentaire. Je reproduis ci-après la version en langue française.

« OAI-PMH

Les métadonnées des publications de l'OIE sont désormais exposées et moissonnables selon le protocole OAI-PMH et le standard de description de métadonnées Dublin Core. Pour plus d'informations sur le protocole OAI-PMH : [www.openarchives.org](http://www.openarchives.org)

L'entrepôt OAI-PMH de l'OIE est disponible à l'adresse :

<http://doc.oie.int:8080/oai/provider>

Exemple de requêtes :

### **Identify**

Informations sur le serveur OAI-PMH.

Exemple : <http://doc.oie.int:8080/oai/provider?verb=Identify>

### **ListMetadataFormats**

Récupérer les formats de métadonnées disponibles (Dublin Core : oai\_dc).

Exemple : <http://doc.oie.int:8080/oai/provider?verb=ListMetadataFormats>

Paramètres :

- identifier (optionnel)

### **ListSets**

Récupérer la structure des différents ensembles ou sets.

Exemple : <http://doc.oie.int:8080/oai/provider?verb=ListSets>

Paramètres :

- resumptionToken (exclusif)

### **ListIdentifiers**

Forme abrégée de ListRecords, pour ne récupérer que les identifiants des jeux de données.

Exemple : <http://doc.oie.int:8080/oai/provider?verb=ListIdentifiers>

### **ListRecords**

Moissonner les enregistrements.

Exemple :

[http://doc.oie.int:8080/oai/provider?verb=ListRecords&metadataPrefix=oai\\_dc](http://doc.oie.int:8080/oai/provider?verb=ListRecords&metadataPrefix=oai_dc)

Paramètres :

- from (optionnel, format AAAA-MM-JJ%)

- until (optionnel, format AAAA-MM-JJ%)
- metadataPrefix (obligatoire)
- set (optionnel)
- resumptionToken (exclusif)

## GetRecord

Récupérer les métadonnées d'un enregistrement individuel.

Exemple :

`http://doc.oie.int:8080/oai/provider?verb=GetRecord&metadataPrefix=oai_dc&identifier=oai:doc.oie.int:dc_32316`

Paramètres :

- identifier (obligatoire)
- metadataPrefix (obligatoire) »

## Figure 8 : Exemple de notice récupérable en OAI-PMH

```
<OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
<responseDate>2015-10-14T10:34:22Z</responseDate>
<request verb="GetRecord" identifier="oai:doc.oie.int:dc_32316"
metadataPrefix="oai_dc">http://doc.oie.int:8080/oai/provider</request>
<GetRecord>
<record>
<header>
<identifier>oai:doc.oie.int:dc_32316</identifier>
<timestamp>2015-10-12T11:13:31Z</timestamp>
```

```

<setSpec>rst</setSpec>
</header>
<metadata>
<oai_dc:dc xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-
instance" xmlns:dcterms="http://purl.org/dc/terms/" xmlns:dc="http://purl.org/dc/elements/1.1/"
xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/ http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
<dc:title>
International standards: the World Organisation for Animal Health Terrestrial Animal Health Code
</dc:title>
<dcterms:alternative xml:lang="fr">
Normes internationales : le Code sanitaire pour les animaux terrestres de l'Organisation mondiale de la santé animale
</dcterms:alternative>
<dcterms:alternative xml:lang="es">
Normas internacionales: el Código sanitario para los animales terrestres de la Organización Mundial de Sanidad Animal
</dcterms:alternative>
<dc:creator>THIERMANN A.</dc:creator>
<dc:contributor>OIE</dc:contributor>
<dc:date>2015-04-01</dc:date>
<dcterms:issued>2015-04-01</dcterms:issued>
<dc:language>eng</dc:language>
<dc:description>
pp. 277-281. 3 ref. This paper provides a description of the international standards contained in the Terrestrial Animal Health
Code of the World Organisation for Animal Health (OIE) that relate to the prevention and control of vector-borne diseases. It
identifies the rights and obligations of OIE Member Countries regarding the notification of animal disease occurrences, as well as
the recommendations to be followed for a safe and efficient international trade of animals and their products.
</dc:description>
<dcterms:abstract xml:lang="en">
This paper provides a description of the international standards contained in the Terrestrial Animal Health Code of the World
Organisation for Animal Health (OIE) that relate to the prevention and control of vector-borne diseases. It identifies the rights and
obligations of OIE Member Countries regarding the notification of animal disease occurrences, as well as the recommendations to
be followed for a safe and efficient international trade of animals and their products.
</dcterms:abstract>
<dcterms:abstract xml:lang="fr">
L'auteur décrit les normes internationales relatives à la prévention et au contrôle des maladies à transmission vectorielle
contenues dans le Code sanitaire pour les animaux terrestres de l'Organisation mondiale de la santé animale (OIE). Il expose les
droits et les obligations des Pays Membres de l'OIE concernant la notification des cas de maladies animales, ainsi que les
recommandations à suivre en vue de garantir l'innocuité et l'efficacité des échanges internationaux d'animaux et de produits
d'origine animale.
</dcterms:abstract>
<dcterms:abstract xml:lang="es">
El autor describe las normas internacionales contenidas en el Código sanitario para los animales terrestres de la Organización
Mundial de Sanidad Animal (OIE) que guardan relación con la prevención y el control de enfermedades transmitidas por vectores.
Asimismo, expone los derechos y obligaciones de los Países Miembros de la OIE en cuanto a notificación de casos de
enfermedades animales, así como las recomendaciones que conviene seguir para mayor seguridad y eficacia del comercio
internacional de animales y productos de origen animal.

```

```

</dcterms:abstract>
<dc:subject>ANALYSE DU RISQUE</dc:subject>
<dc:subject>CERTIFICAT VETERINAIRE</dc:subject>
<dc:subject>CODE SANITAIRE POUR LES ANIMAUX TERRESTRES</dc:subject>
<dc:subject>ECHANGE</dc:subject>
<dc:subject>SANTE ANIMALE</dc:subject>
<dc:subject>SURVEILLANCE</dc:subject>
<dc:subject>NOTIFICATION DES MALADIES</dc:subject>
<dc:subject>OIE (ORGANISATION MONDIALE DE LA SANTE ANIMALE)</dc:subject>
<dc:subject>NORME INTERNATIONALE</dc:subject>
<dc:subject>PAYS MEMBRE</dc:subject>
<dc:format>application/pdf</dc:format>
<dc:type>Text</dc:type>
<dc:type>Article</dc:type>
<dc:identifiant>0253-1933</dc:identifiant>
<dc:identifiant>http://217.19.49.201:8080/dyn/portal/index.seam?page=alo&aloId=32316</dc:identifiant>
<dcterms:isPartOf>Rev. sci. tech. Off. int. Epiz., 34-1. New developments in major vector-borne diseases. Part One: An
overview</dcterms:isPartOf>
<dc:rights>© OIE (World Organisation for Animal Health), 2015</dc:rights>
</oai_dc:dc>
</metadata>
</record>
</GetRecord>
</OAI-PMH>

```

### 3.5 Fichiers KBART

KBART (*Knowledge Bases and Related Tools*) est un guide de bonnes pratiques édité par la National Information Standards Organization (NISO, organisme américain à but non lucratif dédié aux normes dans le domaine de l'édition, des bibliothèques et de l'accès à l'information) [9, NISO].

Son objectif est d'améliorer les échanges de métadonnées relatives aux ressources électroniques avec les bases de connaissances et les résolveurs de liens, notamment afin d'augmenter l'efficacité du processus de construction de liens OpenURL.

Associé à un entrepôt OAI-PMH, KBART permet de répondre à la totalité des recommandations en matières d'échanges de métadonnées avec les bases de connaissance sur les publications académiques numériques. L'intérêt pour l'OIE était donc de pouvoir compléter l'ouverture de l'entrepôt OAI-PMH par des informations plus ciblées, à destination des bases de connaissance, pour un référencement correct des publications numériques et des droits de consultation associés.

Concrètement, il s'agit de faire parvenir au moins une fois par an des fichiers KBART aux bases de connaissance publiques (BACON, qui transmet à KB+ et GoKB) ou commerciales (Summon, Primo, Ebsco DT, WorldCat DT). J'ai donc écrit un premier fichier test que j'ai soumis à l'ABES, dans le cadre du projet de base de connaissance BACON. Conformément au guide de bonnes pratiques, le fichier était au format .txt avec 25 champs bibliographiques listant les publications numériques de l'OIE ainsi que les droits associés (embargo, accès au texte intégral, etc.). Le retour de l'ABES fut très positif, la qualité du fichier fourni étant jugée excellente.

Au-delà d'une remarque sur l'encodage nécessaire du fichier en UTF-8, les points soulevés par mon interlocuteur étaient :

- proposition de l'écriture de trois fichiers KBART différents : l'un pour les périodiques et intitulé « OIE\_Global\_AllJournals\_AAAA-MM- JJ.txt » ; l'autre pour les livres électroniques, intitulé « OIE\_Global\_AllEbooks\_AAAA-MM- JJ.txt ». Enfin, une masterlist comprenant l'ensemble des métadonnées contenues dans chacun des deux fichiers précédents et intitulée « OIE\_Global\_AllTitles\_AAAA-MM-JJ.txt » (le nommage des fichiers étant strictement défini dans le guide de bonnes pratiques, il conviendra de substituer AAAA-MM-JJ par la date de production des fichiers) ;
- flou sur la définition de la période d'embargo (pour les acteurs de l'information, il s'agit de la période pendant laquelle un article de revue numérique ne peut être

accessible librement. Embargo ne devrait donc pas être utilisé pour désigner le décalage de publication entre une version imprimée et une version numérique ;

- application d'une licence CC0 (< <http://creativecommons.org/> >) pour les métadonnées contenues dans le fichier, condition indispensable à leur réutilisation ;
- nommage des URL relatifs aux titres périodiques (dans la colonne « title\_url ») : il est recommandé de créer des pages dédiées au périodique, avec un nom court (voir l'exemple d'Oxford University Press, infra 3.1.1. Architecture et navigation au sein du site internet de l'OIE).

Dans l'attente d'une décision sur les deux derniers points, je n'ai pas fait parvenir de fichier à Bacon.

### **3.6 Positionnement sur le libre accès**

Aujourd'hui, la *Revue scientifique et technique* est disponible en libre accès sur le site de l'OIE. Plus précisément, tous les articles de chaque numéro sont accessibles dans leur intégralité, au format PDF, dès leur mise en ligne. Cependant, cette mise en ligne intervient quelques mois après la parution de l'édition imprimée de la revue. Ce délai – très court – est appelé en interne « embargo ». Pourtant, selon la définition courante appliquée à ce terme dans le monde du libre accès, la revue n'a pas de période d'embargo dans sa version numérique : l'accès au texte intégral des articles numérisés n'est pas restreint pendant une période donnée (par abonnement ou paiement à la carte) mais est immédiatement ouvert à tous. Bien qu'il reste à sortir de l'ambiguïté liée au terme « embargo », on considérera donc que la revue répond au modèle hybride : diffusion imprimée payante, libre accès à la version numérique.

À l'heure où l'ouverture des données s'inscrit dans notre quotidien et où les vertus de l'ouverture font l'objet d'un large consensus, le libre accès peut sembler particulièrement attrayant. Pour autant, s'engager dans l'open access dans sa version la plus stricte impliquerait des choix structurels importants pour l'OIE, dont il convient de bien peser les avantages et les inconvénients.

En septembre 2015, une rencontre avec l'équipe de la revue *Parasite* (< <http://www.parasite-journal.org/> >) a ainsi permis de présenter des pistes de réflexions.

Par ailleurs, j'ai mis à profit l'étude des conditions de référencement sur le site DOAJ (il s'agissait d'une demande exprimée au sein de l'OIE au début du stage) pour illustrer le positionnement actuel de la *Revue scientifique et technique* au regard d'une certaine définition de l'*open access*.

Comme je l'ai indiqué précédemment (voir infra 2.3.2.6. DOAJ), le Directory of Open Access Journals < [www.doaj.org](http://www.doaj.org) > est un annuaire des revues académiques à comité de lecture disponibles en libre accès, principalement autour du modèle dit de la voie dorée.

Pour être référencé dans DOAJ, il convient de renseigner un formulaire en ligne de 58 questions (un fichier est disponible pour préparer la saisie en ligne : < [https://docs.google.com/spreadsheets/d/15zwwg2nXHigzAGIEluLGrB7w-LdbHYGGu\\_ZkFd3ueRU/edit?pli=1#gid=0](https://docs.google.com/spreadsheets/d/15zwwg2nXHigzAGIEluLGrB7w-LdbHYGGu_ZkFd3ueRU/edit?pli=1#gid=0) >, consulté le 15 septembre 2015), portant sur les caractéristiques générales de la revue, la qualité et la transparence du comité de rédaction et des processus de relecture, le libre accès, les licences applicables, le copyright et les autorisations/cessions de droit. Ce formulaire est transmis à quatre membres bénévoles qui évalueront l'éligibilité de la revue.

Un délai de quatre mois est annoncé pour le traitement de chaque demande de référencement. En cas de rejet, un délai incompressible de six mois est exigé avant



toute nouvelle demande (un an si le rejet est lié à des informations erronées dans le formulaire). Dans les trois mois qui suivent l'acceptation de sa demande, l'éditeur doit transmettre les métadonnées de la revue et, s'il le souhaite, celles de tous les articles disponibles (fichier XML conforme à un schéma .xsd établi par DOAJ).

Ci-après sont présentés les points clés, impératifs ou fortement recommandés par DOAJ, à prendre en compte avant de soumettre une demande de référencement. La politique de l'OIE sur ces questions était en cours de discussion à la fin de mon stage ; une mise en application immédiate de mes recommandations (signalées ci-après en caractères italiques soulignés) aurait donc été prématurée.

### **Engagement en faveur de l'open access**

Toute revue académique qui demande son référencement dans DOAJ doit fournir un accès immédiat au texte intégral des articles, sans période d'embargo.

Un engagement clair en faveur de l'open access, en conformité à la définition donnée par la Budapest Open Access Initiative (< <http://www.budapestopenaccessinitiative.org/> >, consulté le 15 septembre 2015) doit être présent sur le site de la revue. Voici, à titre d'exemple, le modèle recommandé d'engagement :

*"This is an open access journal which means that all content is freely available without charge to the user or his/her institution. Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the articles in this journal without asking prior permission from the publisher or the author. This is in accordance with the BOAI definition of open access."*

(Je propose la traduction suivante : « Cette revue est en libre accès, ce qui implique que l'ensemble de son contenu est librement disponible, à titre gratuit pour l'utilisateur ou son institution. Les utilisateurs peuvent lire, télécharger, copier, distribuer, imprimer, fouiller ou lier aux textes intégraux des articles de cette revue sans avoir à solliciter l'avis préalable de l'éditeur ou de l'auteur. Tout ce qui précède est en accord avec la définition du libre accès dite BOAI. »)

Notons que le modèle recommandé par DOAJ semble éloigné des choix actuels de l'OIE. D'autres formulations sont possibles, l'essentiel étant d'indiquer un positionnement clair en faveur de l'open access. Dans tous les cas, il pourra être utile que l'OIE explicite le décalage entre la version imprimée et la version numérique. DOAJ refuse tout embargo pour l'accès aux versions numériques ; l'OIE est donc en conformité mais devra veiller à éviter toute confusion sur le terme embargo.

## **Site Web dédié**

Chaque revue doit avoir son propre site Web.

La structure du site doit favoriser la navigation des lecteurs et l'indexation par les moteurs de recherche. Ainsi, la page d'accueil du site de la revue doit être claire et concise, et doit proposer une navigation immédiate vers le numéro en cours, les archives ou numéros précédents, un outil de recherche, le comité de rédaction et les contacts actualisés. Elle doit proposer également des liens vers les pages d'informations essentielles du journal (objectifs et champs d'application de la revue, comité de rédaction, instructions aux auteurs, système de contrôle qualité, engagements sur le libre accès, éthique éditoriale et plagiat, etc.). Celles-ci doivent être disponibles sur le site et non sur des sites tiers, et doivent démontrer des engagements éthiques élevés.

Tout contenu lié à la revue doit être présent sur ce site et non dispersé sur plusieurs sites.

Le contenu de la revue doit être disponible article par article avec :

- un URL unique par article ;
- un seul PDF par article.

Chaque article doit être daté.

L'utilisation de DOI est fortement recommandée, de même qu'une structure par volume/numéro/page.

La revue doit avoir un ISSN pour sa version imprimée et un autre pour la version numérique.

Il est recommandé d'utiliser un service d'archivage pour la préservation à long terme des documents numériques.

*Je renvoie ici à mes recommandations présentées plus haut (voir infra 3.1.1. Architecture et navigation au sein du site internet de l'OIE).*

### **Comité de rédaction / éthique éditoriale**

Chaque article doit suivre un processus qualité rigoureux avant publication. Ce processus doit être explicitement décrit sur le site Web de la revue.

DOAJ recommande d'intégrer les points suivants dans les instructions aux auteurs :

- Une feuille de style ;
- Une description du processus qualité ;
- Une information sur le copyright de l'article après publication ;

- Un engagement contre le plagiat.

Un engagement contre le plagiat doit idéalement être consultable sur le site et mentionner le ou les logiciels de détection utilisés.

Les revues sont encouragées à prendre des engagements forts en matière d'éthique éditoriale.

À cette fin, signalons l'existence du COPE (Committee on Publication Ethics), créé en 1997, qui regroupe plus de 9000 éditeurs de publications académiques qui s'engagent à suivre son Code of Conduct and Best Practice Guidelines for Journal Editors (< <http://publicationethics.org/resources> >, consulté le 15 septembre 2015). L'adhésion annuelle au COPE coûte environ GBP150.

À titre d'exemple, la revue Parasite affiche sur son site internet l'engagement suivant :

"Ethics for animal experiments, medical studies and scholarly publishing should be irreproachable. Parasite adheres to COPE (Committee of Publishing Ethics)."

Le recours à un logiciel de détection de plagiat est fortement recommandé par DOAJ.

CrossRef propose un tel service avec CrossCheck, qui permet d'interroger un immense corpus de publications académiques dont une majorité d'articles payants habituellement inaccessibles au contrôle des éditeurs.

## **Copyright, cessions de droits, creative commons**

DOAJ exige une très grande clarté sur les droits attachés aux articles. L'application d'une licence Creative Commons est recommandée, notamment parce que cela montre aux auteurs les limites imposées pour les travaux dérivés.

Il est recommandé de communiquer clairement la politique de la revue en matière de dépôt d'articles (auto-archivage par les auteurs).

Signalons que plusieurs sites existent pour permettre aux auteurs et partenaires de visualiser rapidement la politique d'une revue OA en matière de dépôt d'articles et de réutilisation des publications :

- Publications internationales : <http://www.sherpa.ac.uk/romeo/>
- Publications francophones : <http://heloise.ccsd.cnrs.fr/>
- Publications espagnoles ou hispanophones : <http://www.accesoabierto.net/dulcinea/>

Le premier de ces sites (Sherpa Romeo) est notamment utilisé comme référence par le réseau social professionnel Research Gate pour déterminer les conditions d'affichage des publications ajoutées par les utilisateurs.

**Figure 9 : Exemple d'une politique de dépôt d'articles sur le site Héloïse**

Détail de la politique d'archivage d'une revue

Revue	Journal de Chirurgie Viscérale		
Support	Soumis	Validé	Version Editeur
Archives ouvertes	✓	✓	✗
Site personnel	✓	✓	✗
Intranet	✓	✓	✗

## Figures 10 et 11 : Exemples d'une politique de dépôt d'articles sur le site Sherpa Romeo

Journal:	<a href="#">Journal of Virology</a> (ISSN: 0022-538X, EISSN: 1098-5514)
RoMEO:	This is a <a href="#">RoMEO green</a> journal
Paid OA:	A paid open access option is <b>available</b> for this journal.
Author's Pre-print:	✓ author can archive pre-print (ie pre-refereeing)
Author's Post-print:	✓ author can archive post-print (ie final draft post-refereeing)
Publisher's Version/PDF:	✓ author can archive publisher's version/PDF
General Conditions:	<ul style="list-style-type: none"> <li>• Author's pre-print on recognised non profit pre-print archives</li> <li>• Author's post-print on funder's repositories, institutional repository or subject-based repositories</li> <li>• Non-commercial</li> <li>• Publisher's version/PDF may be used</li> <li>• Publisher's version/PDF may be used on author's personal website or employers website</li> <li>• Recommended that author's post-prints submitted to PubMed or institutional repositories are made available 6 months after publication</li> </ul>
Mandated OA:	(Awaiting information)
Paid Open Access:	<a href="#">Optional Open Access</a>
Notes:	<ul style="list-style-type: none"> <li>• Publisher last contacted on 21/05/2015</li> </ul>
Copyright:	<a href="#">Authors Rights - Instructions for Authors (Example)</a>
Updated:	04-Jun-2015 - <a href="#">Suggest an update for this record</a>
Link to this page:	<a href="http://www.sherpa.ac.uk/romeo/issn/0022-538X/">http://www.sherpa.ac.uk/romeo/issn/0022-538X/</a>
Published by:	<a href="#">American Society for Microbiology</a> - <i>Green Policies in RoMEO</i>
<p>This summary is for the journal's <i>default</i> policies, and changes or exceptions can often be negotiated by authors.  <i>All information is correct to the best of our knowledge but should not be relied upon for legal advice.</i></p>	

Journal:	<a href="#">Parasite</a> (ISSN: 1252-607X, EISSN: 1776-1042) [started 2013]
RoMEO:	This is a <a href="#">RoMEO green</a> journal
Paid OA:	A paid open access option is <b>available</b> for this journal.
Listed in:	<a href="#">DOAJ</a> as an open access journal
Author's Pre-print:	✓ author can archive pre-print (ie pre-refereeing)
Author's Post-print:	✓ author can archive post-print (ie final draft post-refereeing)
Publisher's Version/PDF:	✓ author can archive publisher's version/PDF
General Conditions:	<ul style="list-style-type: none"> <li>• On author's personal website or institutional website or OAI compliant website</li> <li>• Some journals require an embargo for deposit in funder's designated repositories (see journal)</li> <li>• Publisher's version/PDF may be used (see journal)</li> <li>• Must link to publisher version</li> <li>• Publisher copyright and source must be acknowledged</li> <li>• Non-commercial</li> </ul>
Mandated OA:	(Awaiting information)
Paid Open Access:	<a href="#">Open Access Option</a>
Copyright:	<a href="#">Example Policy - Preprint servers / ArXiv - Example Copyright Policy</a>
Updated:	27-Feb-2013 - <a href="#">Suggest an update for this record</a>
Link to this page:	<a href="http://www.sherpa.ac.uk/romeo/issn/1252-607X/">http://www.sherpa.ac.uk/romeo/issn/1252-607X/</a>
Published by:	<a href="#">EDP Sciences</a> (according to DOAJ) - <i>Green Policies in RoMEO</i>
<p>This summary is for the journal's <i>default</i> policies, and changes or exceptions can often be negotiated by authors.  <i>All information is correct to the best of our knowledge but should not be relied upon for legal advice.</i></p>	

**Figure 12 : Codes couleur des politiques de dépôt d'articles sur le site Sherpa Romeo**

ROMEIO colour	Archiving policy
<a href="#">green</a>	can archive pre-print <i>and</i> post-print or publisher's version/PDF
<a href="#">blue</a>	can archive post-print (ie final draft post-refereeing) or publisher's version/PDF
<a href="#">yellow</a>	can archive pre-print (ie pre-refereeing)
<a href="#">white</a>	archiving not formally supported

**Figure 13 : Exemple de fiche éditeur sur le site DOAJ**

## Parasite

1252-607X (Print); 1776-1042 (Online)



[Homepage](#)  
**Publisher:** EDP Sciences  
**Society/Institution:** Societe française de parasitologie  
**Country of publisher:** France  
**Platform/Host/Aggregator:** EDP Sciences  
**Date added to DOAJ:** 19 Feb 2015

**LCC Subject Category:** Medicine: Internal medicine: Infectious and parasitic diseases  
**Publisher's keywords:** acarology, helminthology and protistology, epidemiology, parasites, parasitology  
**Language of fulltext:** English  
**Full-text formats available:** PDF, HTML, ePUB

### PUBLICATION CHARGES

Article Processing Charges (APCs): **Yes**. 900EUR  
 Submission Charges: **No**.  
 Waiver policy for charges? **No**.

### EDITORIAL INFORMATION

[Blind peer review](#)  
[Editorial Board](#)  
[Aims and scope](#)  
[Instructions for authors](#)  
 Time From Submission to Publication: 8 weeks

[Less](#)

### OPEN ACCESS & LICENSING

Type of License: CC BY  
[License terms](#)  
[Open Access Statement](#)  
 Year Open Access Content Began: 2013  
 Does the author retain copyright? **Yes**  
 Does the author retain publishing rights? **Yes**

### BEST PRACTICE POLICIES

Permanent article identifier: DOI  
 Checks for plagiarism: **Yes**  
 Content digitally archived in: [PMC/Europe PMC/PMC Canada](#)  
 Deposit policy registered in: Héloïse

# Conclusion



Le dynamisme de certains groupes de travail interprofessionnels, la convergence des bonnes pratiques, la normalisation de nombreux formats et protocoles ont considérablement simplifié l'interopérabilité des métadonnées des revues académiques. Les questions techniques liées au référencement reposent au final sur des outils standards bien documentés. Par rapport aux objectifs de départ du projet, j'ai eu la bonne surprise de pouvoir mettre en œuvre rapidement des solutions (KBART, OAI-PMH, fichiers XSLT) qui, de prime abord, m'avaient semblé exiger des moyens et savoirs techniques importants.

La concentration des acteurs – qu'il s'agisse du poids de Google Scholar, de la domination d'une poignée de bases d'indexation ou des quatre principaux acteurs du marché des outils de découverte – permet d'obtenir un effet de leviers considérable en se focalisant sur quelques objectifs clés.

Échanger les métadonnées des articles de revue académique pour améliorer leur visibilité et leur « trouvabilité » est une opération dont la complexité réside moins dans l'aspect technique que dans la structuration et la qualité des données. C'est là tout l'intérêt de la gouvernance des métadonnées mise en œuvre à l'OIE : le portail documentaire regroupe dans une base unique l'ensemble des informations bibliographiques sur les publications.

Mais avoir des données de qualité et les outils pour les exposer ne constitue qu'une étape. Il est essentiel de maintenir un contact fort avec les acteurs qui réutiliseront ces métadonnées, notamment les bases d'indexation et les outils de découverte, pour garantir l'adéquation des flux mis en place avec les besoins nouveaux ou à venir de ces partenaires. D'autre part, dans un contexte en constante évolution, il s'avère nécessaire de maintenir une veille sur les normes et bonnes pratiques pour l'interopérabilité, le référencement naturel, les réseaux sociaux professionnels, les politiques de libre accès, etc. Enfin, et surtout, il convient de garder à l'esprit que la qualité des données – comme leur pérennité – implique un engagement continu, dont le coût n'est pas nul.

Le projet que j'ai conduit nécessitait une phase exploratoire du logiciel documentaire Alexandria, afin d'exposer au mieux les métadonnées de la *Revue scientifique et technique*. Le recours au langage XML, à des protocoles tels que l'OAI-PMH, à des formats tels que le Dublin Core, à des identifiants pérennes tels que les DOI était adapté aux besoins du projet mais correspondait aussi aux standards du web sémantique. Si le chemin reste encore long avant de pouvoir s'engager dans la voie du web de données, il est intéressant de constater qu'une grande partie du travail effectué peut constituer la première étape d'une telle démarche, qui pourra s'appuyer à l'avenir sur de nombreuses fonctionnalités déjà présentes dans Alexandria.

Pour aller plus loin encore, nous pouvons par exemple noter l'intérêt que représenterait l'attribution d'identifiants uniques aux auteurs de la revue. Cela permettrait d'éviter les incertitudes en cas d'homonymie, tout en favorisant les liens et rebonds à partir d'un auteur. Si les ISNI (*International Standard Name Identifier* < <http://www.isni.org> >) semblent moins pertinents car attribués assez tardivement par les bibliothèques, les identifiants ORCID (*Open Researcher and Contributor ID* < <http://orcid.org> >) paraissent particulièrement bien adaptés pour les auteurs de l'OIE, majoritairement des professionnels en activité.

# **Bibliographie**

La présente bibliographie est de type analytique : chaque référence est augmentée d'un commentaire présentant l'apport théorique ou pratique qui m'a été utile pour ce projet.

Elle est conforme aux normes Z44-005 de décembre 1987 et NF ISO 690-2 de février 1998. Elle est arrêtée au 15 octobre 2015 et elle est organisée en 7 sections.

## Généralités

[1] BERMÈS, Emmanuelle, ISAAC, Antoine et POUPEAU, Gauthier, 2013. *Le Web sémantique en bibliothèque*. Paris : Electre : Éd. du Cercle de la Librairie. ISBN 978-2-7654-1417-9.

*Avec une approche très pratique, cet ouvrage complet sur le web sémantique en bibliothèque donne les clés essentielles pour structurer les données afin de les rendre interopérables. Après une première partie décrivant les technologies en œuvre, une seconde partie est dédiée par la présentation de cas pratiques. Ce livre m'a permis de privilégier les solutions qui préparent le mieux la base documentaire de l'OIE aux défis du web de données.*

[2] BOURDENET, Philippe, 2013. *L'espace documentaire en restructuration : l'évolution des services des bibliothèques universitaires*. [en ligne]. Thèse de doctorat. Conservatoire national des arts et métiers - CNAM. [Consulté le 16 septembre 2015]. Disponible à l'adresse : <https://tel.archives-ouvertes.fr/tel-00932683/>

*Cette thèse de doctorat analyse les mutations des espaces documentaires en bibliothèques universitaires et les nouveaux modèles de données favorisant les liaisons avec l'univers de la connaissance. L'auteur analyse notamment les moyens nécessaires pour une évolution vers le web sémantique, pour valoriser les collections.*

*L'exhaustivité des descriptions fournies en fait une référence essentielle sur l'interopérabilité des métadonnées bibliographiques.*

[3] CLAERR, Thierry, WESTEEL, Isabelle, ORY-LAVOLLÉE, Bruno et GEORGES, Nicolas, 2013. *Manuel de constitution de bibliothèques numériques*. Paris : Éd. du Cercle de la librairie. ISBN 978-2-7654-1413-1.

*Bien que le sujet de l'ouvrage soit connexe au projet, j'y ai fréquemment eu recours en raison de la très grande clarté des présentations des principaux concepts liées aux métadonnées bibliographiques. Sur l'interopérabilité, l'OAI-PMH, les DOI, le Dublin Core et beaucoup d'autres points traités ici, ce manuel a constitué une propédeutique idéale.*

[4] WARE, Mark et WABE, Michael, 2015. *The STM Report. An overview of scientific and scholarly journal publishing*. [en ligne]. International Association of Scientific, Technical and Medical Publishers. [Consulté le 16 septembre 2015]. Disponible à l'adresse : [http://www.stm-assoc.org/2015\\_02\\_20\\_STM\\_Report\\_2015.pdf](http://www.stm-assoc.org/2015_02_20_STM_Report_2015.pdf)

*Ce rapport sur les revues académiques commandé par la International Association of STM Publishers < <http://www.stm-assoc.org> > dresse un état-de-l'art de l'édition mondiale en sciences, techniques et médecine. Compilant un grand nombre de sources, il regroupe les données les plus récentes sur le secteur. La qualité et la diversité des informations fournies en font un document essentiel à tout professionnel intéressé par les revues académiques.*

## Base de connaissance et outils de découverte

[5] BREEDING, Marshall et NATIONAL INFORMATION STANDARDS ORGANIZATION (U.S.), 2015. *The future of library resource discovery: a white paper commissioned by the NISO Discovery to Delivery (D2D) Topic Committee* [en ligne]. [Consulté le 16 septembre 2015]. ISBN 9781937522414 1937522415. Disponible à l'adresse :

[http://www.niso.org/apps/group\\_public/download.php/14487/future\\_library\\_resource\\_discovery.pdf](http://www.niso.org/apps/group_public/download.php/14487/future_library_resource_discovery.pdf)

*Libre blanc issu du groupe de travail Open Discovery Initiative de l'agence de normalisation américaine NISO, ce document explore le futur des outils de découverte en bibliothèque. Sa lecture m'a éclairé sur les enjeux liés à l'évolution probable de ces outils. Ce document m'a aussi permis de fonder mes recommandations sur les métadonnées destinées aux bases d'indexation et aux bases de connaissance.*

[6] CICCONE, Karen et VICKERY, John, 2015. Summon, EBSCO Discovery Service, and Google Scholar: A Comparison of Search Performance Using User Queries. *Evidence Based Library and Information Practice*. 6 mars 2015. Vol. 10, n° 1, pp. 34-49.

*Cet article compare les performances de deux outils de découverte commerciaux (EBSCO DiscoveryService et Summon/ProQuest) par rapport à Google Scholar (échantillon de 183 recherches réelles d'utilisateurs de toute discipline, sélectionnées au hasard à partir de 2013). Les auteurs n'ont pas conclu à une différence majeure entre les outils d'EBSCO et de ProQuest, mais ont noté la supériorité de Google Scholar. Cet article m'a permis de mieux appréhender le rôle des outils de découverte dans les stratégies de recherche d'information en université, et de rester conscient de l'efficacité croissante de Google Scholar.*

[7] INGER, Simon et GARDNER, Tracy, 2013. Library technology in content discovery – evidence from a large-scale reader survey: Based on a paper presented at the 36th UKSG Annual Conference, Bournemouth, April 2013. *Insights: the UKSG journal*. 1 juillet 2013. Vol. 26, n° 2, pp. 120-127. DOI 10.1629/2048-7754.75.

*Cet article présente les résultats d'une étude de grande envergure sur la recherche d'information à partir des outils disponibles en bibliothèques universitaires (bases d'indexation, outils de découverte, moteurs de recherche, partages de références bibliographiques, etc.). C'est après la lecture de cet article que j'ai eu l'idée d'exprimer les besoins d'interopérabilité des métadonnées en m'appuyant sur les stratégies de navigation pour la recherche d'information.*

[8] LEVINE-CLARK, Michael, MCDONALD, John et PRICE, Jason S, 2014. The effect of discovery systems on online journal usage: a longitudinal study: Updated from a paper presented at the 37th UKSG Annual Conference, Harrogate, April 2014. *Insights: the UKSG journal*. 1 novembre 2014. Vol. 27, n° 3, pp. 249-256. DOI 10.1629/2048-7754.153.

*Cet article étudie le rôle et l'efficacité des outils de découverte dans l'identification et la consultation des articles de revues. Sa lecture m'a conforté dans ma volonté de travailler sur la présence des publications de l'OIE dans ces outils de découverte.*

[9] NISO RP-9-2014. *Knowledge Bases and Related Tools (KBART) Recommended Practice*, 2014. [en ligne]. National Information Standards Organization. [Consulté le 16 septembre 2015]. Disponible à l'adresse : [http://www.niso.org/apps/group\\_public/project/details.php?project\\_id=122](http://www.niso.org/apps/group_public/project/details.php?project_id=122)

*Les fichiers KBART permettent l'échange standardisé de métadonnées sur les ressources numériques, entre les éditeurs et les bases de connaissance. Ce guide de bonnes pratiques regroupe l'ensemble des informations nécessaires à la création d'un tel fichier. Je l'ai utilisé pour produire le fichier KBART relatif aux ressources numériques de l'OIE.*

## DOI, pérennité

[10] January 2015 DOI Outage: Followup Report | CrossTech, [sans date]. [en ligne]. [Consulté le 4 juin 2015]. Disponible à l'adresse : <http://crosstech.crossref.org/2015/03/january-2015-doi-outage-followup-report.html>

*Ce billet de blog fait suite à la panne géante qui a frappé le système DOI en janvier 2015. Au-delà d'une description des causes de l'incident (l'absence de renouvellement du nom de domaine doi.org), le billet analyse de façon originale la question de la pérennité et propose quelques pistes pour l'avenir du système DOI. Sa lecture m'a également permis de valider que la panne était liée au système dans son ensemble et non à la seule agence CrossRef.*

[11] KLEIN, Martin, VAN DE SOMPEL, Herbert, SANDERSON, Robert, SHANKAR, Harihar, BALAKIREVA, Lyudmila, ZHOU, Ke et TOBIN, Richard, 2014. Scholarly Context Not Found: One in Five Articles Suffers from Reference Rot. *PLoS ONE*. 26 décembre 2014. Vol. 9, n° 12, pp. e115253. DOI 10.1371/journal.pone.0115253.

*Cet article présente une étude sur les liens morts (contenu déplacé ou modifié) dans les références des articles académiques, qui, dès lors, ne peuvent être replacés dans le contexte qui a entouré leur rédaction. La grande quantité de données utilisées permet d'étayer tout travail de sensibilisation à la persistance des informations, par exemple au travers de la pérennité des métadonnées.*



## Dublin Core et OAI-PMH

[12] SIREAU, Christine, 2012. La valorisation sur Internet du patrimoine scientifique : étude des besoins. Le cas de la médiathèque scientifique de l'Institut Pasteur [en ligne]. other. Institut national des techniques de la documentation du CNAM. [Consulté le 16 septembre 2015]. Disponible à l'adresse : [http://memsic.ccsd.cnrs.fr/mem\\_00803701/document](http://memsic.ccsd.cnrs.fr/mem_00803701/document)

*Ce mémoire de l'INTD présente la mise en place d'une bibliothèque numérique pour valoriser les documents patrimoniaux de la Médiathèque scientifique de l'Institut Pasteur. Sa lecture m'a particulièrement intéressé pour l'analyse des difficultés de mise en place d'un entrepôt OAI-PMH en lien avec un logiciel documentaire.*

[13] *Guide d'utilisation du Dublin Core à la Bibliothèque nationale de France*, [sans date]. [en ligne]. Bibliothèque nationale de France. [Consulté le 16 septembre 2015]. Disponible à l'adresse : [http://www.bnf.fr/documents/guide\\_dublin\\_core\\_bnf\\_2008.pdf](http://www.bnf.fr/documents/guide_dublin_core_bnf_2008.pdf)

*Ce guide, assez court, présente l'utilisation du langage Dublin Core non qualifié (15 éléments de base) à la BNF. Sa lecture m'a été précieuse lorsque, pour créer l'entrepôt OAI-PMH, j'ai dû préparer les données et aligner les étiquettes des champs de données du logiciel documentaire avec leur équivalence en Dublin Core non qualifié. Les informations fournies sur certains formats de données (langues, dates, etc.) m'ont resservi lorsque j'ai par la suite travaillé sur le Dublin Core qualifié.*

## Libre accès

[14] BOMSEL, Olivier, 2014. Copyright et économie des revues académiques. *Revue européenne des sciences sociales*. 6 mai 2014. N° 52-1, pp. 93-110. DOI 10.4000/ress.2673.

*Bien que centré sur les impacts du libre accès en sciences humaines et sociales, cet article m'a particulièrement éclairé sur la transformation des modèles économiques pour l'édition de revue académique. Les impacts de la numérisation et du libre accès sur les marques éditoriales font l'objet d'analyses originales. Cet article éclaire les différences entre l'univers scientifique, technique et médical et celui des sciences humaines.*

[15] JAFFRIN, Ivan et PARISOT, Thomas, 2014. La place des revues dans la communication scientifique en régime de libre accès: Retour sur le débat français en SHS. *Revue européenne des sciences sociales*. 6 mai 2014. N° 52-1, pp. 9-35. DOI 10.4000/ress.2637.

*Réagissant à une polémique sur l'accès aux informations scientifiques, cet article explore ce que pourrait être l'avenir de la communication scientifique en régime de libre accès, à l'aune de de deux grands paradigmes (« légal-bibliométrique » vs « collaboratif-en-réseau »). Bien que centré sur les sciences humaines et sociales, il est particulièrement éclairant sur l'analyse du basculement possible d'un filtre à l'entrée (l'accès à la publication via la validation en amont par les pairs) vers un filtre à la sortie, par le biais des outils bibliométriques et collaboratifs. Après lecture de cet article, j'ai veillé à formuler des recommandations sur les réseaux sociaux académiques.*

## Référencement naturel

[16] NISO RP-16-2013, *PIE-J: Presentation & Identification of E-Journals*, 2013. [en ligne]. National Information Standards Organization. [Consulté le 15 octobre 2015]. Disponible à l'adresse :

[http://www.niso.org/apps/group\\_public/download.php/10368/rp-16-2013\\_pie-j.pdf](http://www.niso.org/apps/group_public/download.php/10368/rp-16-2013_pie-j.pdf)

*Au travers des nombreux exemples de sites Web de revues académiques, dont la structuration des informations fait l'objet d'une analyse avancée, ce guide de bonnes pratiques démontre la convergence des pratiques considérées comme meilleures, car plus efficaces en terme de visibilité. J'ai abondamment puisé dans ce guide lorsque j'ai étudié l'organisation des pages de la Revue scientifique et technique sur le site de l'OIE.*

## Sciences vétérinaires

[17] JISC, 2009. Virtual Training Suite: Internet for Veterinary medicine. [en ligne]. 2009. [Consulté le 5 juin 2015]. Disponible à l'adresse :

<http://www.vtstutorials.co.uk/ws//tracking/launchcontent.aspx?cv=E57F6038-520D-45DA-A541-FD5086B2333E&e=A0000&c=2EDD079D-1428-4539-9E1F-FD729E13E45A&SID=2aebc079-156e-4852-b1b1-19f3efc712c1>

*Destiné aux étudiants britanniques, ce mini-site à vocation pédagogique présente les bases d'une recherche d'information de qualité en médecine vétérinaire. L'un des exemples porte sur les zoonoses, domaine dans lequel l'OIE fait référence. Je recommande ce site à toute personne souhaitant pouvoir s'orienter rapidement parmi les ressources numériques disponibles en médecine vétérinaires, comme cela fut mon cas (notamment pour identifier les ressources d'acteurs comparables à l'OIE).*

# **Annexes**

## Annexe 1 Citations croisées avec CrossRef

J'ai souhaité mesurer les contraintes induites par l'obligation faite par CrossRef d'ajouter les DOI dans les références des articles.

### Méthode :

Test de l'outil manuel disponible à l'adresse < <http://www.crossref.org/simpleTextQuery> > (les éditeurs affiliés à CrossRef ont accès à une API pour le traitement par lots, plus rapide).

Exemple avec copier/coller dans l'outil CrossRef des vingt premières références de l'article :

J. de la Fuente, K.M. Kocan & M. Contreras, *Prevention and control strategies for ticks and pathogen transmission*, Rev. Sci. Tech. Off. Int. Epiz., 2015, 34 (1), 249-264

1. Dennis D.T. & Piesman J.F. (1991). – Tick-borne diseases of humans (J.L. Goodman, D.T. Dennis & D.E. Sonenshine, eds). ASM Press, Washington, DC, 3–11.

2. De la Fuente J. (2003). – The fossil record and the origin of ticks (Acari: Parasitiformes: Ixodida). *Experim. Appl. Acarol.*, 29 (3–4), 331–344.

3. Sonenshine D.E. (1991). – Introduction. In *Biology of ticks* (D.E. Sonenshine, ed.), Vol. I. Oxford University Press, New York, 3–12.

4. Barker S.C. & Murrell A. (2008). – Systematics and evolution of ticks with a list of valid genus and species names. In *Ticks: biology, disease and control* (A.S. Bowman & P.A. Nuttall, eds). Cambridge University Press, Cambridge, 1–39.

5. Sonenshine D.E., Lane R.S. & Nicholson W.L. (2002). – Ticks (Ixodida). In *Medical and veterinary entomology* (G. Mullen & L. Durden, eds). Academic Press, New York, 517–558.

6. De la Fuente J., Estrada-Peña A., Venzal J.M., Kocan K.M. & Sonenshine D.E. (2008). – Overview: ticks as vectors of pathogens that cause disease in humans and animals. *Front. Biosci.*, 13, 6938–6946.
7. Korch G.W. Jr (1991). – Geographic dissemination of tick-borne zoonoses. In *Biology of ticks* (D.E. Sonenshine, ed.), Vol. I. Oxford University Press, New York, 139–197.
8. Jonejan F. & Uilenberg G. (2004). – The global importance of ticks. *Parasitol.*, 129, S3–S14.
9. De Castro J.J. (1997). – Sustainable tick and tick-borne disease control in livestock improvement in developing countries. *Vet Parasitol.*, 71 (2–3), 77–97.
10. Alim M.A., Das S., Roy K., Masuduzzaman M., Sikder S., Hassan M.M., Siddiki A.Z. & Hossain M.A. (2012). – Prevalence of hemoprotozoan diseases in cattle population of Chittagong division, Bangladesh. *Pakistan Vet. J.*, 32 (2), 221–224.
11. Valente P.P., Amorim J.M., Castilho R.O., Leite R.C. & Ribeiro M.F. (2014). – In vitro acaricidal efficacy of plant extracts from Brazilian flora and isolated substances against *Rhipicephalus microplus* (Acari: Ixodidae). *Parasitol. Res.*, 113 (1), 17–23.
12. Munderloh U.G., Jauron S.D. & Kurti T.J. (2005). – The tick: a different kind of host for human pathogens. In *Tick-borne diseases of humans* (J.L. Goodman, D.T. Dennis & D.E. Sonenshine, eds). ASM Press, Washington, DC, 37–64.
13. Smith T. & Kilbourne F.L. (1893). – Investigations into the nature, causation, and prevention of Texas or southern cattle fever. *Bur. Anim. Ind. Bull.*, 1, 177–304.
14. Parola P. & Raoult D. (2001). – Ticks and tick-borne bacterial diseases in humans: an emerging infectious threat. *Clin. Infect. Dis.*, 32 (6), 897–928.
15. Lowes R. (2014). – Discovery of Bourbon virus raises many questions. *Medscape Med. News*, 24 December. Available at: [www.medscape.com/viewarticle/837229](http://www.medscape.com/viewarticle/837229) (accessed on 22 January 2015).
16. Nicholson W., Sonenshine D.E., Lane R.S. & Uilenberg G. (2009). – Ticks (Ixodida). In *Tick-borne diseases of humans* (L.A. Durden & G. Mullen, eds). ASM Press, Washington, DC, 150–163.
17. Sonenshine D.E. & Roe R.M. (eds) (2014). – *The biology of ticks*, 2nd Ed., Vol. I. Oxford University Press, New York, 560 pp.
18. Sutherst R. (2001). – The vulnerability of animal and human health to parasites under global change. *Int. J. Parasitol.*, 31 (9), 933–948.
19. Camill P. (2010). – Global change. *Nat. Educ. Knowl.*, 2, 49.
20. Estrada-Peña A., Ayllón N. & de la Fuente J. (2012). – Impact of climate trends on tick-borne pathogen transmission. *Front. Physiol.*, 3, 64. doi:10.3389/fphys.2012.00064.

## Analyse du résultat :

On obtient une page de résultats reprenant les références chargées, complétées de leurs identifiants PubMed (3 cas) et de DOI (11 cas). Le résultat est visible ci-dessous.

1. Dennis D.T. & Piesman J.F. (1991). – Tick-borne diseases of humans (J.L. Goodman, D.T. Dennis & D.E. Sonenshine, eds). ASM Press, Washington, DC, 3–11.

[PMid:2051923](#)

2. De la Fuente J. (2003). – The fossil record and the origin of ticks (Acari: Parasitiformes: Ixodida). *Experim. Appl. Acarol.*, 29 (3–4), 331–344.

<http://dx.doi.org/10.1023/A:1025824702816>

3. Sonenshine D.E. (1991). – Introduction. In *Biology of ticks* (D.E. Sonenshine, ed.), Vol. I. Oxford University Press, New York, 3–12.

4. Barker S.C. & Murrell A. (2008). – Systematics and evolution of ticks with a list of valid genus and species names. In *Ticks: biology, disease and control* (A.S. Bowman & P.A. Nuttall, eds). Cambridge University Press, Cambridge, 1–39.

<http://dx.doi.org/10.1017/CBO9780511551802.002>

5. Sonenshine D.E., Lane R.S. & Nicholson W.L. (2002). – Ticks (Ixodida). In *Medical and veterinary entomology* (G. Mullen & L. Durden, eds). Academic Press, New York, 517–558.

<http://dx.doi.org/10.1016/b978-012510451-7/50026-8>

6. De la Fuente J., Estrada-Peña A., Venzal J.M., Kocan K.M. & Sonenshine D.E. (2008). – Overview: ticks as vectors of pathogens that cause disease in humans and animals. *Front. Biosci.*, 13, 6938–6946.

<http://dx.doi.org/10.2741/3200>

[PMid:18508706](#)

7. Korch G.W. Jr (1991). – Geographic dissemination of tick-borne zoonoses. In *Biology of ticks* (D.E. Sonenshine, ed.), Vol. I. Oxford University Press, New York, 139–197.

8. Jonejan F. & Uilenberg G. (2004). – The global importance of ticks. *Parasitol.*, 129, S3–S14.

<http://dx.doi.org/10.1017/S0031182004005967>

9. De Castro J.J. (1997). – Sustainable tick and tick-borne disease control in livestock improvement in developing countries. *Vet Parasitol.*, 71 (2–3), 77–97.

[http://dx.doi.org/10.1016/S0304-4017\(97\)00033-2](http://dx.doi.org/10.1016/S0304-4017(97)00033-2)

10. Alim M.A., Das S., Roy K., Masuduzzaman M., Sikder S., Hassan M.M., Siddiki A.Z. & Hossain M.A. (2012). – Prevalence of hemoprotozoan diseases in cattle population of Chittagong division, Bangladesh. *Pakistan Vet. J.*, 32 (2), 221–224.

11. Valente P.P., Amorim J.M., Castilho R.O., Leite R.C. & Ribeiro M.F. (2014). – In vitro acaricidal efficacy of plant extracts from Brazilian flora and isolated substances against *Rhipicephalus microplus* (Acari: Ixodidae). *Parasitol. Res.*, 113 (1), 17–23.  
<http://dx.doi.org/10.1007/s00436-013-3670-2>
12. Munderloh U.G., Jauron S.D. & Kurtti T.J. (2005). – The tick: a different kind of host for human pathogens. In *Tick-borne diseases of humans* (J.L. Goodman, D.T. Dennis & D.E. Sonenshine, eds). ASM Press, Washington, DC, 37–64.  
<http://dx.doi.org/10.1128/9781555816490.ch3>
13. Smith T. & Kilbourne F.L. (1893). – Investigations into the nature, causation, and prevention of Texas or southern cattle fever. *Bur. Anim. Ind. Bull.*, 1, 177–304.
14. Parola P. & Raoult D. (2001). – Ticks and tick-borne bacterial diseases in humans: an emerging infectious threat. *Clin. Infect. Dis.*, 32 (6), 897–928.  
<http://dx.doi.org/10.1086/319347>  
PMid:11247714
15. Lowes R. (2014). – Discovery of Bourbon virus raises many questions. *Medscape Med. News*, 24 December. Available at: [www.medscape.com/viewarticle/837229](http://www.medscape.com/viewarticle/837229) (accessed on 22 January 2015).
16. Nicholson W., Sonenshine D.E., Lane R.S. & Uilenberg G. (2009). – Ticks (Ixodida). In *Tick-borne diseases of humans* (L.A. Durden & G. Mullen, eds). ASM Press, Washington, DC, 150–163.
17. Sonenshine D.E. & Roe R.M. (eds) (2014). – *The biology of ticks*, 2nd Ed., Vol. I. Oxford University Press, New York, 560 pp.
18. Sutherst R. (2001). – The vulnerability of animal and human health to parasites under global change. *Int. J. Parasitol.*, 31 (9), 933–948.  
[http://dx.doi.org/10.1016/S0020-7519\(01\)00203-X](http://dx.doi.org/10.1016/S0020-7519(01)00203-X)
19. Camill P. (2010). – Global change. *Nat. Educ. Knowl.*, 2, 49. J. de la Fuente, K.M. Kocan & M. Contreras, Prevention and control strategies for ticks and pathogen transmission, *Rev. Sci. Tech. Off. Int. Epiz.*, 2015, 34 (1), 249-264
20. Estrada-Pe-a A., Ayllón N. & de la Fuente J. (2012). – Impact of climate trends on tick-borne pathogen transmission. *Front. Physiol.*, 3, 64. doi:10.3389/fphys.2012.00064.  
<http://dx.doi.org/10.3389/fphys.2012.00064>

**En conclusion, cette obligation semble peu contraignante et apporte aux références une grande valeur ajoutée.**