

# Formation initiale des bibliothécaires d'État

# Stratégies pour un dépôt légal des périodiques en ligne

Dossier d'aide à la décision

Maryline DEVIDAL

Sous la direction d'Alix BRUYS

Chef du service Gestion des périodiques – BnF, département du Dépôt légal

Sous la direction de Raphaële MOUREN

Maître de conférences – ENSSIB



#### Remerciements

Je remercie Alix Bruys, Raphaële Mouren, Gildas Illien, Géraldine Raison, Philippe Cantié et Dominique Brégiroux pour le temps qu'ils m'ont consacré, les informations qu'ils m'ont apportées et les conseils qu'ils m'ont prodigués,

Je remercie Catherine Andrieux, Marie Guibert, Jude Talbot et Michèle Devidal, d'avoir eu la gentillesse de relire ces pages et de me faire part de leurs remarques et suggestions, tant sur le fonds que sur la forme.



Résumé:

Dépositaire du dépôt légal imprimé, la BnF est confrontée au report d'une partie

de l'offre éditoriale sur le web, et souhaite développer des stratégies de collecte

de l'édition de périodiques en ligne, notamment dans le cadre de sa nouvelle

mission de dépôt légal de l'internet.

Descripteurs:

Dépôt légal – Périodiques – France

Périodiques électroniques – France

Archives de l'internet – France

Bibliothèque nationale de France

Abstract :

Receiving legal deposit of printed materials, the Bibliothèque nationale de France

has to face the report of the editorial offer on the internet, and is looking for

ways to collect electronic journals, especially in the area of her new assignment :

collecting the internet.

Keywords:

*Legal deposit of books, etc – Periodicals – France* 

*Electronic journals – France* 

*Web archives – France* 

Bibliothèque nationale de France

Droits d'auteurs



(cc) BY-NC-ND

Cette création est mise à disposition selon le Contrat :

Paternité-Pas d'Utilisation Commerciale-Pas de Modification 2.0 France disponible en ligne http://creativecommons.org/licenses/by-nc-nd/2.0/fr/ ou par courrier postal à Creative Commons, 171 Second Street, Suite 300, San

Francisco, California 94105, USA.

# Sommaire

SIGLES ET ABREVIATIONS	9
INTRODUCTION	11
LE DEPOT LEGAL DES IMPRIMES A LA BIBLIOTHEQUE NATIONALE FRANCE	
Organisation	13
Le Département du dépôt légal, DDL	
Missions et organisation	13
Réception et enregistrement	
Catalogage	
Echanges	
Réclamations et prospection	
Le dépôt légal des périodiques  Le Service de gestion des périodiques (GDP)	
La Bibliographie nationale française – périodiques (BFP)	17
Le Centre ISSN France (CIF)	17
Le nouveau modèle économique de la publication en ligne : enjeux	oour
le dépôt légal des périodiques	
L'édition de périodiques et le virage internet	
Exemple de publications hybrides : la presse d'information	
Exemple de publications natives : les webzines	
Exemple de publications migrées : les publications officielles (PO) .	
Analyse d'après les données recueillies par la BnF	
Enjeux pour le dépôt légal des périodiques	
L'APPORT DU DEPOT LEGAL DE L'INTERNET	25
Un nouveau champ d'application pour le département du dépôt lég	
Définition et cadre juridique	
Les modalités de collecte des documents	
Le signalement des données et l'accès aux documents	
Archivage pérenne des collections  La place des périodiques en ligne dans ce dispositif	
La stratégie actuelle de collecte des périodiques en ligne	
Le bilan de ce qui a déjà été collecté dans les Archives de l'internet	
Une réponse adaptée mais qui n'est pas sans soulever de question	
La question de l'exhaustivité de la collecte et de la continuité des coll	
La question du contrôle qualité	36
La question de l'archivage pérenne	
La question du signalement	36
La question de l'accès et de la valorisation	
Enjeux déclinés par origine de la publication	
Titres natifs numériques	
Migrations en ligne	
Ressources hybrides  Lettres d'information électroniques diffusées par courriel et titres	41
disponibles sur abonnement	42



Bilan	42
ELEMENTS DE COMPARAISON	43
Problématiques rencontrées par d'autres services de la BnF	attributaires
du dépôt légal : les livres et les documents audiovisuels	
Le Service de gestion des livres (GDL)	
Le Département de l'audiovisuel (DAV)	
Approches envisagées par des bibliothèques nationales étra	
mettant en œuvre des actions d'archivage du web	
La Bibliothèque royale du Danemark	
La Bibliothèque nationale du Québec	
La Bibliothèque nationale suisse	
SCENARIO A / LA PROBLEMATIQUE DE CONTINUITE DES COL	
SCENARIO A / LA PROBLEMATIQUE DE CONTINUITE DES COL	
La veille et les sélections : mettre en place une stratégie pros	
Veille systématique	
Recueillir l'information auprès de l'éditeur	50
Développer la communication externe	
Aller à la rencontre de l'éditeur : annualisation des chantier	
Exploiter les outils informatiques	
Extractions de Millennium	
Extractions de l'application de veille Prosper	
Exploiter le réseau professionnel interne	
Liens avec GCA Liens avec le CIF	
Liens avec les catalogueurs de BFP et des départements d	
	51
Editer un formulaire de signalement	
Compléter la veille systématique par des chantiers de veille cibi	
Les PO	
La presse d'information : les grands titres de la PQN, de la PQ	
presse magazine	
Les presses universitaires et scientifiques dans le domaine de	
sociales et des sciencesLa presse professionnelle	
La presse associative	
Mettre en œuvre le dépôt légal de substitution pour les publicat	ions dont la
collecte papier pose problème	52
La question des revues disponibles sur abonnement  Mise en œuvre de la collecte par le DLN	
Le contrôle qualité des archives constituées	53
Le signalement : vers un signalement enrichi dans le catalog	ue général
de la BnF	
Signalement de la mort de l'édition imprimée (cas des publication	• ,



(cas des publications hybrides pour lesquelles la collecte papier pose problème)	55
Catalogage de l'édition en ligne et attribution d'un nouvel ISSN	
Accès et valorisation : multiplier les points d'accès à la publication	
Accès à la publication dans le web « vivant »	
Accès à la publication dans les Archives de l'internet	
Projet de refonte du catalogue	
Avantages et inconvenients du scénario A	
SCENARIO B / LA PROBLEMATIQUE D'EXHAUSTIVITE DE LA COLLECTE	
La veille et les sélections : mettre en place une stratégie a la fois	
prospective, concertée et raisonnée	
Inscrire la collecte des périodiques comme l'un des objectifs du DLN	
Constituer et animer le réseau des sélectionneurs	
Viser à l'exhaustivité raisonnée en fixant des priorités et en organisant de	
chantiers ciblés devant alimenter des paniers transverses	
Champs de priorité par typologie de producteurs	
Les PO	
Les revues universitaires et scientifiques	
La presse professionnelle	
La presse associative	
Champs de priorité spécifiques à l'édition en ligne	
Les webzines	
Les revues de sommaires	
La presse d'information en ligne (nationale, régionale, magazine)	.61
Champs de priorité axés sur la collecte des portails ou bases de	64
périodiques	
Les bases d'archives de presseLes kiosques virtuels	
Les portails de revues	
La question des revues disponibles sur abonnement	
La veille	
Les sites fournissant des données sur les éditeurs ou les titres	
Exploiter les outils informatiques	
Extractions du catalogue	
Extractions de Prosper	
Extractions de l'rosper Extraction de l'rosper Extraction extranet	
Mise en œuvre de la collecte par le DLN	
Le contrôle qualité des archives constituées	03 64
Le signalement : adopter une stratégie de signalement a minima	64
Signalement de la mort de l'édition imprimée	
Pas de catalogage de l'édition en ligne	
Politique indépendante du CIF : attribution d'ISSN et catalogage de l'éditi	
en ligne	
L'accès et la valorisation : adopter une stratégie allégée	
L'accès à la publication dans le web « vivant »	
L'accès à la publication dans les Archives de l'internet	
Valorisation auprès de la communauté professionnelle et du public	
Le cas particulier des titres ayant été catalogués par le CIF	. 67
Avantages et inconvénients du scénario b	

BILAN	68
CONCLUSION	69
BIBLIOGRAPHIE	71
TABLE DES ANNEXES	73



# Sigles et abréviations

ABES Agence bibliographique de l'enseignement supérieur

AFNIC Association française pour le nommage internet en

coopération

BFP Service de la Bibliographie nationale française - Périodiques

BnF Bibliothèque nationale de France

CIF Centre ISSN France

CPPAP Commission paritaire des publications et agences de presse

DADVSI (Loi) Loi relative au droit d'auteur et aux droits voisins dans la

société de l'information

DAV Département de l'audiovisuel

DDL Département du dépôt légal

DLN Service du dépôt légal numérique

DSR Direction des services et des réseaux

ETP Equivalent temps plein

GCA Gestion centralisée des acquisitions

GDL Service de gestion des livres

GDP Service de gestion des périodiques

IBN Département de l'information bibliographique et numérique

IIPC International internet preservation consortium

ISSN International Standard Serial Number

PO Publications officielles

PQN Presse quotidienne nationale

PQR Presse quotidienne régionale

SIGB Système intégré de gestion de bibliothèque

SPAR Système de préservation et d'archivage réparti

Sudoc Système universitaire de documentation

URL Uniform ressource locator



#### Introduction

Le dépôt légal est l'obligation pour tout éditeur, imprimeur, producteur, distributeur, importateur, de déposer chaque document qu'il édite, imprime, produit, distribue ou importe en France à la Bibliothèque nationale de France (BnF) ou auprès de l'organisme habilité à recevoir le dépôt, en fonction de la nature du document. Il permet la constitution d'une collection patrimoniale de référence.

La BnF est en charge du dépôt légal des documents imprimés, graphiques, photographiques, sonores, audiovisuels, multimédias, quel que soit leur procédé technique de production, d'édition ou de diffusion, ainsi que des logiciels et bases de données, quelle que soit la nature de leur support. L'Institut national de l'audiovisuel est dépositaire des documents sonores et audiovisuels radio et télé diffusés, tandis que le Centre national du cinéma et de l'image animée est dépositaire de l'ensemble des documents destinés à une première exploitation en salles de spectacles cinématographiques.

Habitués à collecter des documents sur supports, quels qu'ils soient, ces différents organismes sont confrontés depuis un certain nombre d'années à la dématérialisation des contenus et à leur migration vers le web. L'édition en ligne, si elle n'est pas systématique, constitue désormais une part non négligeable de la production éditoriale française, et elle échappe à toute collecte. La BnF souhaite donc trouver des stratégies pour une mise en œuvre de la collecte de ces publications, notamment dans le cadre de sa mission nouvelle de dépôt légal de l'internet.

En termes de dépôt légal des périodiques cette migration des contenus vers le web soulève trois problématiques majeures : celle de l'exhaustivité de la collecte (actuellement toute une frange de l'édition de périodiques échappe au dépôt légal), celle de la continuité des collections (les collections des titres nés imprimés mais migrant vers le web ne sont plus complétées), et celle de la pérennité de l'accès aux documents, puisqu'on le sait le web est un univers très volatile.

Deux types de collecte sont envisageables : le dépôt par l'éditeur des fichiers correspondant à sa publication au format numérique, ou la collecte par la BnF de la publication dans son environnement, le web dynamique. La BnF est en charge depuis 2006 d'une mission nouvelle de dépôt légal : la collecte du domaine français de l'internet. C'est cette nouvelle disposition qui rend possible d'envisager la collecte de la publication dans son environnement dynamique.

L'objectif de ce rapport est de proposer des stratégies de mise en œuvre d'un dépôt légal des périodiques en ligne. Il se fonde pour cela sur un diagnostic quantitatif et qualitatif des publications concernées. Il s'efforce ensuite de cerner les problématiques et les enjeux en cause et de dresser des comparaisons pouvant alimenter la réflexion, à l'intérieur (problématiques et stratégies développées par d'autres services de dépôt légal de la BnF), et à l'extérieur (stratégies mises en œuvre par des bibliothèques menant des projets d'archivage du web).

La commande émane du Service de gestion des périodiques au sein du Département du dépôt légal. Elle consiste en une étude préliminaire devant aboutir à proposer des scénarii à soumettre au Service du dépôt légal numérique, seul juge de ce qu'il est possible ou non de faire en l'état actuel de ses compétences techniques et de ses moyens.

Réfléchir à des stratégies de mise en œuvre du dépôt légal des périodiques en ligne implique de mener une réflexion parallèle sur leur traitement après collecte : signalement, accès et valorisation. Les scénarii proposés engloberont donc des stratégies de collecte et d'archivage, des stratégies de signalement, et des stratégies d'accessibilité et de valorisation des collections constituées.

# Le dépôt légal des imprimés à la Bibliothèque nationale de France

Institué par François ler en 1537<sup>1</sup>, le dépôt légal a pour objectif de collecter l'intégralité de la production éditoriale française, dans le but de la conserver et d'en garantir l'accès aux générations futures. Le dépôt légal est obligatoire et gratuit.

La Bibliothèque nationale de France a pour mission de collecter ces documents, de les signaler, de les conserver et de les communiquer au public, dans le respect du droit de la propriété intellectuelle. Livres, brochures et périodiques sont signalés dans la Bibliographie nationale française et dans le Catalogue général de la BnF<sup>2</sup>. Le dépôt légal est assorti d'une restriction de la consultation sur place, et les collections patrimoniales ainsi constituées sont consultables à la BnF, dans la bibliothèque de recherche, par les usagers accrédités.

Le dépôt légal sur le territoire français métropolitain et d'outre-mer est régi par le Code du patrimoine³ pour les aspects législatifs et par le décret n°93-1429 du 31 décembre 1993⁴ pour les aspects réglementaires. Le Code du patrimoine est entré en vigueur le 20 février 2004. Il remplace la loi n° 92-546 du 20 juin 1992⁵ relative au dépôt légal. Il a notamment été modifié par la loi n° 2006-961 du 1er août 2006⁶ relative au droit d'auteur et aux droits voisins dans la société de l'information. Les articles qui fixent le cadre législatif du dépôt légal figurent dans le "Livre 1er : Dispositions communes à l'ensemble du patrimoine culturel - Titre III : Dépôt légal³" du Code du patrimoine.

### ORGANISATION8

# Le Département du dépôt légal, DDL

#### Missions et organisation

Le Code du patrimoine désigne explicitement la BnF comme étant en charge d'organiser le dépôt légal des imprimés afin de constituer les collections patrimoniales. Le Département du dépôt légal est responsable de la collecte et

<sup>&</sup>lt;sup>8</sup> Voir en annexe 1 les organigrammes de la BnF, du Département du dépôt légal et du Département de l'information bibliographique et numérique



<sup>&</sup>lt;sup>1</sup> Ordonnance de Montpellier

<sup>&</sup>lt;sup>2</sup> http://catalogue.bnf.fr, consulté le 01/09/2010

<sup>3</sup>http://www.legifrance.gouv.fr/affichCode.do;jsessionid=762CB5AE8FFD2C95E816ADEC5AEAC267.tpdjo05v\_2?cidTexte=LEGITEXT000006074236&dateTexte=20100819, consulté le 01/09/2010

<sup>4</sup> http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006082758&dateTexte=20090223, consulté le 01/09/2010

<sup>&</sup>lt;sup>5</sup> <a href="http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006079299&dateTexte=20100902">http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006079299&dateTexte=20100902</a>, consulté le 01/09/2010

<sup>&</sup>lt;sup>6</sup>http://www.legifrance.gouv.fr/html/actualite/actualite\_legislative/decrets\_application/2006-961.htm, consulté le 01/09/2010

 $<sup>\</sup>frac{\text{http://www.legifrance.gouv.fr/affichCode.do?idArticle=LEGIARTI000006845515\&idSectionTA=LEGISCTA00000615993}{4\&\text{cidTexte}=\text{LEGITEXT000006074236\&dateTexte}=20090917},\ consult\'e\ le\ 01/09/2010$ 

du traitement bibliographique des imprimés entrés par dépôt légal, à l'exception des documents spécialisés : cartes et plans, musique imprimée, estampes et photographies, qui sont perçus par les départements spécialisés. La gestion de la collecte implique à la fois de veiller à l'exhaustivité et à la continuité des collections. Le département gère des flux, et n'assure pas la conservation des documents. Cette charge incombe aux départements de collections.

Au sein de la Direction des services et des réseaux (DSR), le DDL est organisé en six services, soit 156 agents au 31 décembre 2009 (146,8 équivalents temps pleins). Les effectifs sont répartis comme suit, en dehors de la direction et du secrétariat général : 33 agents pour le Service de gestion des périodiques, 26 agents pour le Service de gestion des livres, 64 agents pour la Bibliographie nationale française - livres, 9 agents pour la Bibliographie nationale française - périodiques, 17 agents pour le Service des échanges, et 5 agents pour le Service du dépôt légal numérique en charge de la collecte de l'internet.

#### Réception et enregistrement

Cette gestion implique la réception matérielle des documents imprimés livres et périodiques envoyés par la poste ou déposés directement sur place en deux exemplaires<sup>9</sup>.

Dès leur réception, livres et périodiques sont enregistrés dans le Système d'information de la Bibliothèque selon une filière spécifique dédiée au dépôt légal. Conformément aux décrets suscités le dépôt s'accompagne obligatoirement d'une déclaration de dépôt légal en 3 exemplaires. Après vérification de la conformité du dépôt et de la qualité des données saisies dans la déclaration, le département renvoie le premier exemplaire au déposant en guise de récépissé de dépôt légal et archive le second exemplaire. Le troisième exemplaire est mis à la disposition des services de catalogage du département.

La mise en service en 2009 de l'extranet du dépôt légal des éditeurs permet la dématérialisation de cette procédure de déclaration. Les éditeurs de livres ou de périodiques ont désormais la possibilité de faire leurs déclarations en ligne. Les données sont versées directement dans le catalogue permettant ainsi un premier signalement succinct mais rapide. L'extranet produit un récépissé électronique à l'intention des éditeurs et la Bibliothèque est dispensée de l'envoi du formulaire papier.

#### Catalogage

Une fois les entrées enregistrées, les documents imprimés, livres et publications en série sont ensuite pris en charge par les services de bibliographie, afin d'être signalés dans le Catalogue général de la bibliothèque et de paraître dans la Bibliographie nationale française. Ils sont ensuite ventilés dans les départements thématiques de la bibliothèque pour y rejoindre les magasins et enrichir les collections patrimoniales.

<sup>&</sup>lt;sup>9</sup> Le dépôt se fait en un seul exemplaire si le tirage est inférieur à 300 exemplaires DEVIDAL Maryline | FIBE T | Projet professionnel personnel | 2010



-

#### **Echanges**

Le second exemplaire est redistribué à des bibliothèques partenaires, en France exclusivement. Cette redistribution se fait selon des critères très précis ayant abouti à l'élaboration d'une carte documentaire.

Le Service des échanges gère également les échanges de documents 10 avec des partenaires étrangers. Les documents, acquis de façon à refléter la production éditoriale française de niveau scientifique, sont envoyés à l'intention des institutions partenaires en contrepartie des documents reçus de leur part.

#### Réclamations et prospection

Bien que la loi impose à tout producteur de document imprimé de procéder au dépôt légal dès lors que le document est mis en nombre à disposition d'un public, des lacunes dans les dépôts sont constatées et de nombreuses réclamations sont effectuées tant pour les livres que pour les publications en série.

### Le dépôt légal des périodiques

#### Le Service de gestion des périodiques (GDP)

Au sein du DDL, GDP assure la collecte et l'enregistrement des périodiques au titre du dépôt légal éditeur et importateur ainsi qu'au titre du dépôt légal imprimeur pour la région lle-de-France. Il veille à la complétude de la collecte, à la continuité des collections et à l'exhaustivité des dépôts, ce pour quoi il s'est doté d'une cellule de veille éditoriale. Il participe également au traitement des nouveaux titres avant leur transmission au service de catalogage.

Le service enregistre 41 437 titres vivants au 1<sup>er</sup> janvier 2010. Cela représente un accroissement de 1,6 % par rapport à l'année 2008. En termes de volumes cela représente 343 270 dépôts en 2009, soit 690 000 fascicules/an, soit en moyenne 2 800 fascicules/jour. La création de 3 692 nouveaux titres<sup>11</sup> est en baisse de 6 % par rapport à 2008, compensée par la diminution des disparitions de titres. Sur 2 571 cessations enregistrées, 249 ont migré vers une publication au format numérique, soit 10 %.

Les principales activités du service sont le tri et le bulletinage, au cours desquelles sont réalisées les opérations de réception, dépouillement, compostage, enregistrement des numéros dans le module bulletinage du SIGB<sup>12</sup> Millennium depuis décembre 2005, report des cotes et ventilation des exemplaires dans les départements de conservation. C'est également à ce moment que sont assurées les opérations de maintenance des cardex informatiques, de repérage des nouveaux titres et des changements de titres, et les relations avec les déposants. Une équipe des circuits est chargée de maintenir la fluidité de la filière en élaborant des consignes et des procédures.

<sup>&</sup>lt;sup>12</sup> Système intégré de gestion de bibliothèque



<sup>&</sup>lt;sup>10</sup> Livres et périodiques uniquement

<sup>&</sup>lt;sup>11</sup> Dont 341 entrés par veille au cours de l'année 2009

Le module de bulletinage de Millennium incluant des fonctionnalités de prévision et de gestion des retards, il permet de gérer les réclamations de façon semi-automatique. En 2009 48 286 réclamations ont ainsi été envoyées, soit environ 4 000 par mois. Deux chantiers de traitement des retards ayant occasionné cinq réclamations restées sans suite 13, ont été menés en 2009 et en 2010. Ces chantiers ont permis, suite à des contacts téléphoniques personnalisés avec les éditeurs, soit de récupérer les numéros manquants, soit de mettre à jour les cardex 14.

Afin d'assurer la complétude et l'exhaustivité de la collecte, le service s'est doté d'une cellule de veille éditoriale chargée de la prospection des nouveaux titres et des réclamations sur des titres déclarés morts ou en suspens alors que la publication a repris. L'objectif affiché est de tendre à l'exhaustivité mais face à l'ampleur de la tâche et au regard des moyens alloués, le service s'est donné un objectif qualitatif en privilégiant ce qui est susceptible d'enrichir les collections. Les sources d'information sont variées. La première consiste traditionnellement à effectuer un suivi régulier de la presse et des portails spécialisés (Tarif Média<sup>15</sup>, Titres info<sup>16</sup>, la Correspondance de la presse<sup>17</sup>, Journaux.fr<sup>18</sup>, Presstalis.fr<sup>19</sup>, Presseedition.fr<sup>20</sup>, etc). Par ailleurs, une veille sectorielle peut être un moyen efficace de tendre à l'exhaustivité sur certains types de publications. C'est le cas notamment pour les publications officielles, les bulletins municipaux, les publications des sociétés savantes, les publications universitaires et autres publications scientifiques. Le réseau professionnel joue également un grand rôle dans le recueil d'informations, en particulier le réseau des bibliothèques habilitées à recevoir le dépôt légal imprimeur en région. Les BDLI fournissent régulièrement à la cellule de veille des copies des pages de couvertures des nouveaux titres qu'elles reçoivent en dépôt, permettant ainsi d'opérer un contrôle croisé. Le Centre ISSN France est également un important pourvoyeur d'informations. Il retransmet les données concernant des titres non déposés ayant fait l'objet d'une demande de numérotation, ainsi que les dossiers de création ou de réouverture de notices, à la demande de l'Agence bibliographique de l'enseignement supérieur (ABES). En 2009, ce sont 761 prospections et 77 réclamations de veille qui ont été lancées. Le contrôle croisé dépôt légal éditeur/imprimeur a été à l'origine de 55,19 % des prospections réalisées, devenant ainsi la première source d'information de la veille. La proportion du nombre de titres entrés par la veille sur l'ensemble des nouveaux titres enregistrés était de 9,2 % ce qui représente une part significative et atteste de l'importance de l'activité.

GDP dispose d'une position privilégiée pour étudier les composantes de l'offre éditoriale de périodiques française. C'est l'un des objectifs inscrits au contrat de performance 2009-2013 du DDL.

<sup>13</sup> Chantiers « récla 5 »

DEVIDAL Maryline | FIBE T | Projet professionnel personnel | 2010

<sup>&</sup>lt;sup>14</sup> Report des informations de cessation, suspension, passage en ligne, changement d'adresse, etc

http://www.tarifmedia.com, consulté le 01/09/2010

http://www.presstalis.fr/actu/frameTitresInfos.htm, consulté le 01/09/2010

Pas de version en ligne

http://www.journaux.fr, consulté le 01/09/2010

http://www.presstalis.fr, consulté le 01/09/2010

http://www.presseedition.fr, consulté le 01/09/2010

# La Bibliographie nationale française – périodiques (BFP)

Service interne du DDL, la BFP a pour mission de réaliser le signalement bibliographique de référence des nouveaux titres<sup>21</sup> de périodiques édités ou diffusés en France, reçus par dépôt légal. Il produit la partie « Publications en série » de la Bibliographie nationale française publiée en ligne. En revanche, la mise à jour des données bibliographiques concernant les cessations et les résurrections est entièrement prise en charge par GDP.

Les deux services, GDP et BFP, sont en relation très étroite. Repérés et mis de côté lors du bulletinage, les nouveaux titres ou changements de titres sont traités hebdomadairement par l'équipe des circuits de GDP, qui statue notamment sur la filière de traitement et rédige une pré-notice dans Millennium avant de transmettre les fascicules à la Bibliographie nationale. Une fois catalogués les fascicules repartent à GDP où les notices de bulletinage et les cardex sont créés et le fascicule est bulletiné. Des réclamations rétrospectives peuvent être lancées à ce moment si le numéro déposé n'est pas le premier.

Le délai médian de création des notices était de trois semaines seulement en 2009. Il importe en effet que le délai de catalogage des périodiques soit le plus rapide possible en raison des flux importants supportés au niveau des entrées. Ceci est particulièrement vrai pour les titres à périodicité rapide, et ce de façon à éviter tout engorgement au moment du bulletinage des fascicules.

#### Le Centre ISSN France (CIF)

Le CIF est un service du Département de l'information bibliographique et numérique (IBN), et non du DDL. Membre du Réseau de l'ISSN, il a pour mission de coordonner et de gérer la numérotation des publications en série éditées en France et d'informer les éditeurs.

L'ISSN<sup>22</sup> est attribué gratuitement et sans formalité aux revues, magazines, collections, etc, à réception du premier numéro ou volume envoyé à la BnF au titre du dépôt légal. L'attribution d'un ISSN n'est pas systématique et dépend de critères précis. Elle est faite par la BFP au moment du catalogage des nouveaux titres arrivés dans les filières d'entrée du dépôt légal. Le CIF est amené lui aussi à attribuer des numéros ISSN à des éditeurs qui en font la demande avant dépôt légal, à la demande de l'ABES<sup>23</sup> pour le Sudoc<sup>24</sup>, ainsi qu'à des éditeurs en ligne. Il emploie 2,8 ETP<sup>25</sup> pour le catalogage des titres auxquels il attribue un ISSN.

L'ISSN est une clé d'accès et un moyen de contrôle qui facilite la gestion automatisée des documents. Il est notamment utilisé dans les applications informatiques pour la mise à jour des fichiers, l'établissement de liens entre fichiers différents, ainsi que la recherche et l'échange de données. L'ISSN intéresse à ce titre un nombre important d'acteurs : éditeurs, diffuseurs, agences

(cc)) BY-NC-ND



<sup>&</sup>lt;sup>21</sup> Créations et changements de titre

<sup>&</sup>lt;sup>22</sup> International standard serial number

<sup>&</sup>lt;sup>23</sup> Agence bibliographique de l'enseignement supérieur

<sup>&</sup>lt;sup>24</sup> Système universitaire de documentation

<sup>&</sup>lt;sup>25</sup> Equivalent temps plein

d'abonnement, libraires, bibliothécaires et documentalistes. Il est un repère nécessaire dans l'univers complexe des publications en série<sup>26</sup>.

Le CIF entretient des rapports étroits avec GDP. Il lui transmet notamment les données concernant des demandes de numérotation pour des titres n'ayant fait l'objet d'aucun dépôt, ainsi que les demandes de créations ou de réouvertures faites par l'ABES.

# LE NOUVEAU MODELE ECONOMIQUE DE LA PUBLICATION EN LIGNE : ENJEUX POUR LE DEPOT LEGAL DES PERIODIQUES

### L'édition de périodiques et le virage internet

La chute des ventes et des recettes publicitaires, la concurrence des gratuits, des blogs d'actualités et des listes de discussion, ainsi que les opportunités de diffusion élargie à moindre coût et d'enrichissement des contenus offertes par les nouvelles technologies numériques, encouragent les éditeurs de périodiques à se tourner vers l'édition numérique en ligne. Sans parler d'une massification de l'édition en ligne ou d'un abandon du modèle économique traditionnel de la presse, on constate un report non négligeable d'une partie de l'offre sur internet.

On constate cependant à la lecture des *Chiffres clés de la presse 2008-2009*<sup>27</sup> que la presse imprimée n'accuse pas pour autant un recul dramatique. En 2007, le chiffre d'affaires de la presse s'élève à 10,70 milliards d'euros, soit un niveau supérieur à 2006, la baisse de la publicité commerciale (- 1,1 %) étant compensée par la hausse des ventes par abonnement (+ 1,7 %) et des ventes au numéro (+ 1,7 %). Les investissements de la presse sont principalement composés des investissements magazines avec 30,9% de l'ensemble des investissements. Suivent ensuite les gratuits avec 24,3 %, la presse quotidienne régionale (23,6 %), la presse spécialisée (11,3 %), la presse quotidienne nationale (6,9 %) et la presse hebdomadaire régionale avec 3% des investissements publicitaires de la presse. Il apparaît donc que les deux modèles peuvent coexister, l'équilibre restant à trouver.

Suite aux états généraux de la presse écrite lancés le 2 octobre 2008 et afin d'encourager le virage internet, un statut d'éditeur de presse en ligne a été institué par le décret n°2009-1340 du 29 octobre 2009 pris pour application de l'article 1er de la loi n° 86-897 du 1er août 1986 portant réforme du régime juridique de la presse <sup>28</sup>. Une aide de l'Etat au développement de la presse en ligne a également été décidée. Institué par le décret n°2009-1379 du 11 novembre 2009<sup>29</sup>, le fonds d'aide au développement des services de presse en ligne <sup>30</sup> succède au fonds d'aide au développement des services en ligne des

<sup>30</sup> fonds SPEL



<sup>&</sup>lt;sup>26</sup> titres éphémères, relations ambiguës entre titres

<sup>&</sup>lt;sup>27</sup> Les chiffres-clés 2008-2009, TarifMédia, http://www.tarifmedia.com/static/services/infos/chiffres\_cles.pdf, consulté le 01/09/2010

<sup>&</sup>lt;sup>28</sup>http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000021214264&dateTexte=&categorieLien=id, consulté le 01/09/2010

<sup>&</sup>lt;sup>29</sup>http://legifrance.gouv.fr/affichTexte.do:jsessionid=1CD2D373ECE6031A3483978CCE21DE06.tpdjo09v\_2?cidTexte= JORFTEXT000021259498&dateTexte=20100909, consulté le 01/09/2010

entreprises de presse, créé par le décret n°2004-1313 du 26 novembre 2004<sup>31</sup>. En 2009, ce sont 15,35 milliards d'euros qui ont ainsi été versés aux éditeurs en ligne.

Alain Jacquesson et Alexis Rivier dressent une typologie des périodiques en ligne dans l'ouvrage *Bibliothèques et documents numériques*<sup>32</sup>: revues scientifiques, revues de sommaires, lettres d'information, listes de discussion, revues mouvantes<sup>33</sup>, web'zines<sup>34</sup>, éditions en ligne des quotidiens et hebdomadaires imprimés, rapports institutionnels, etc... A cette typologie il faut ajouter les critères d'existence ou non d'un contrôle scientifique, le mode d'accès (en ligne avec ou sans abonnement ou par courriel), la gratuité ou la facturation, le format des fichiers (pdf, html, etc), l'existence d'un contenu multimédia ou web 2.0, ou encore l'existence d'une fonction d'autoarchivage avec revue de sommaires.

Il est parfois difficile de distinguer un périodique imprimé d'une collection éditoriale ou d'un recueil. Il est encore plus difficile d'identifier un périodique en ligne au sein du magma hétéroclite de documents que l'on rencontre sur le web. Surtout, les frontières entre les publications numériques sont très perméables et les définitions bibliothéconomiques traditionnelles volent en éclats. Si les périodiques migrés conservent en général le modèle traditionnel basé sur la permanence d'une page de titre, d'un sommaire et d'une périodicité et se réduisent finalement à un fichier pdf au sein d'un site hébergeur, les périodiques natifs en ligne adoptent un modèle beaucoup plus souple où le site lui-même est le périodique, où le sommaire laisse place aux rubriques et où les articles sont susceptibles d'être mis à jour en permanence. C'est le modèle adopté notamment par les webzines, mais aussi par les quotidiens et magazines d'information dans leur version en ligne. Il est donc primordial d'établir des critères pouvant amener à distinguer un périodique d'un site web. Un document rédigé par la BFP en 2004 propose une définition et une typologie suivant quatre critères principaux : l'existence d'un titre constant et d'un directeur de publication, l'existence d'une parution à intervalles réguliers, la vocation informationnelle de la publication, la mise à disposition du public en général ou de catégories de public. Toutefois, le critère de l'existence d'une parution à intervalles réguliers ne semble pas très pertinent rapporté à l'univers souple de l'internet, et nous proposons de l'abandonner.

On peut distinguer trois types de modèles éditoriaux : les publications hybrides pour lesquelles l'édition imprimée continue à paraître en marge de l'édition en ligne, les publications natives en ligne, et les publications migrées. En marge existent des publications qui ne sont pas à proprement parler des périodiques en ligne puisqu'elles sont diffusées par courriel : les lettres d'information électroniques.

<sup>&</sup>lt;sup>34</sup> Informations brutes et décousues, sans structure formelle en général



<sup>31</sup> http://www.legifrance.gouv.fr/affichTexte.do;jsessionid=E448933AD19FE48458F5D28283ABFA75.tpdjo14v\_2?cidTexte=LEGITEXT000005913858&dateTexte=20100322&categorieLien=id, consulté le 01/09/2010

<sup>&</sup>lt;sup>32</sup> JACQUESSON, Alain, RIVIER, Alexis, *Bibliothèques et documents numériques : concepts, composantes, techniques et enjeux*, Paris, Ed. du Cercle de la librairie, 2005 (Collection Bibliothèques)

<sup>&</sup>lt;sup>33</sup> Seul un numéro de la revue est disponible en ligne, et les articles sont mis à jour en permanence

#### Exemple de publications hybrides : la presse d'information

La presse d'information a été une des premières à envisager le tournant numérique. Dès le milieu des années 1990, sentant pointer la concurrence internet, les grands quotidiens nationaux et régionaux se sont implantés sur le web. En 2007 d'après une étude citée dans le Rapport Tessier<sup>35</sup> la quasi-totalité de la presse française est présente sur internet. D'après la même étude, seulement 11% des 37 titres de presse magazine étudiés y sont représentés, et un tiers d'entre eux ne propose en ligne que le sommaire du magazine papier.

Les versions en ligne ont tout d'abord connu des succès mitigés, principalement parce qu'elles ne parvenaient pas à trouver leur identité par rapport à l'édition imprimée. Par ailleurs, elles constituaient une concurrence dangereuse pour cette dernière. L'offre a alors évolué en incluant des contenus de plus en plus distincts et originaux de l'édition imprimée, notamment multimédias, ainsi que des fonctionnalités du web 2.0, voire un contenu rédactionnel propre. On constate pour tous ces titres une continuité de l'existence de l'édition imprimée parallèlement à l'édition en ligne. Le modèle le plus répandu est celui de la gratuité d'accès aux contenus du jour, couplée à un accès payant aux archives. C'est le modèle proposé notamment par le journal Le Monde 36. A côté de ces publications de presse imprimées qui se tournent vers le numérique, on trouve également des titres ayant adopté le modèle inverse : nés numériques, ils lancent leur version imprimée après plusieurs années d'existence en ligne. Le site d'actualités Rue 8937 par exemple, né en ligne le 6 mai 2007, a publié son premier mensuel imprimé le 1<sup>er</sup> juillet 2010.

#### **Exemple de publications natives : les webzines**

Les webzines sont des publications spécifiques de l'environnement internet. Né numérique, un webzine 38 est un site web regroupant des articles originaux. Réalisé par des amateurs bénévoles ou par des salariés, il n'a pas d'existence sur support imprimé. On parle de « webzine interactif » lorsque les visiteurs ont la possibilité d'ajouter leurs commentaires aux articles existants.

Comme les fanzines, les webzines sont souvent thématiques. Les thèmes abordés tournent souvent autour de sujets peu traités par les médias traditionnels : bande dessinée (*Numo*<sup>39</sup>), jeu de rôle (*Jeux d'ombres*<sup>40</sup>), internet (*Xrings*<sup>41</sup>), musique et cinéma (*Discordance*<sup>42</sup>), actualité locale (*Lyon webzine*<sup>43</sup>), etc. Peu de critères objectifs différencient un webzine amateur d'un site personnel : le nombre de contributeurs, leur passion, la qualité technique de la réalisation sont comparables. Couramment, on attend d'un webzine une ambition rédactionnelle, c'est à dire un contenu original et une parution

http://lyon.france-webzine.com, consulté le 01/09/2010 DEVIDAL Maryline | FIBE T | Projet professionnel personnel | 2010



<sup>&</sup>lt;sup>35</sup> TESSIER. Marc, *La presse au défi du numérique : rapport au Ministre de la culture et de la communication*, février 2007, disponible en ligne http://www.culture.gouv.fr/culture/actualites/rapports/tessier/rapport-fev2007.pdf (consulté le 01/09/2010

<sup>36</sup> http://www.lemonde.fr, consulté le 01/09/2010

http://www.rue89.com, consulté le 01/09/2010

<sup>&</sup>lt;sup>38</sup> Mot-valise formé de la contraction de « web » et « magazine »

<sup>39</sup> http://www.numo.fr, consulté le 01/09/2010

<sup>40</sup> http://jeux.dombres.free.fr, consulté le 01/09/2010

http://www.xrings.net, consulté le 01/09/2010

http://www.discordance.fr, consulté le 01/09/2010

relativement régulière. On constate toutefois une grande instabilité sur ce type de publications, l'expérience montrant que la durée de vie d'un titre est généralement assez courte.

# Exemple de publications migrées : les publications officielles (PO)

La dématérialisation de l'administration inscrite dans la politique de modernisation de l'Etat, ainsi que les objectifs de transparence et de performance de l'action publique inscrits dans la loi organique n°2001-692 du 1 août 2001 relative aux lois de finances<sup>44</sup>, conduisent les services publics, qu'ils soient centraux, déconcentrés ou territoriaux, ainsi que les établissements publics, à substituer le mode de diffusion électronique au mode de diffusion imprimé traditionnel dans leurs diverses publications. Les recueils d'actes administratifs sont tout particulièrement concernés par cette migration massive. De même, le Ministère des finances édite pratiquement exclusivement en ligne dorénavant.

Il n'existe pas en réalité de modèle homogène concernant les publications officielles en ligne. Un certain nombre d'administrations ou d'institutions continuent à éditer sous forme papier parallèlement à l'édition en ligne, comme par exemple La Documentation française<sup>45</sup>. Le *Journal officiel* est encore édité sur papier mais existe aussi en ligne. Il existe une version authentifiée depuis le 2 juin 2004 sur le site du Journal officiel<sup>46</sup>, et une version non authentifiée depuis 1990 sur le site de Légifrance<sup>47</sup>. Les versions imprimées et en ligne ne sont pas identiques.

L'Agence du patrimoine immatériel de l'Etat est chargée de faciliter l'accès et la réutilisation de l'information publique. Parmi les actions engagées, la conception d'un portail unique de l'information publique est particulièrement intéressante. Ce projet, devant offrir au public un outil au service de l'intérêt général et de la croissance économique, a été intégré dans le plan gouvernemental « France numérique 2012 ».

# Analyse d'après les données recueillies par la BnF

Deux services de la BnF sont susceptibles de recueillir des données permettant d'analyser l'édition en ligne : GDP et le CIF.

Compte tenu de ses faibles ressources en personnel, le CIF ne tient pas de statistiques par typologie de documents mais des statistiques globales : 339 ISSN ont été attribués à des ressources en ligne en 2008, soit 6% du nombre total d'ISSN attribués. Ce chiffre était de 425 en 2009 et le taux progressait également sensiblement. Environ 3 000 ISSN ont été attribués à ce jour à des ressources en ligne. Ces chiffres sont toutefois à tempérer, le CIF attribuant des numéros ISSN à toutes ressources continues en ligne, c'est-à-dire des périodiques mais aussi des bases de données, des blogs, des sites, etc.

<sup>47</sup> http://www.legifrance.gouv.fr/initRechJO.do, consulté le 01/09/2010



<sup>44</sup> http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000005631294&dateTexte=vig, consulté le 01/09/2010

<sup>45</sup> http://www.ladocumentationfrancaise.fr, consulté le 01/09/2010

http://www.journal-officiel.gouv.fr, consulté le 01/09/2010

GDP de son côté, en raison de sa mission relative à la collecte de l'édition imprimée, tient des statistiques uniquement sur les cessations au profit d'une migration en ligne. Il s'avère donc qu'il n'existe pas aujourd'hui à la BnF de recueil de données statistiques concernant l'édition de périodiques native en ligne ou hybride, ce qui ne nous permet pas d'en dessiner les contours.

Les seules données fines exploitables sont celles recueillies par GDP sur les publications migrées. GDP dispose en effet d'une situation privilégiée pour étudier les titres imprimés ayant migré sur le net. L'éditeur n'informe pas systématiquement de la cessation de l'édition imprimée au profit d'une édition en ligne, mais cela arrive assez régulièrement, notamment lorsque les dépôts sont traités par son service abonnements. Dans ce cas GDP reçoit un courrier comme n'importe quel abonné. D'autre part, dans la mesure où la tête de collection est passée par le service des entrées, celui-ci dispose d'informations permettant le suivi du titre. Ainsi, lorsque l'éditeur n'a pas pris la peine d'avertir de la migration de sa publication, GDP constatant par le biais du module de bulletinage de Millennium que les fascicules n'arrivent plus, adresse un courrier de réclamation à l'éditeur. C'est par ce biais, le plus souvent, que l'on obtient l'information de cessation auprès de l'éditeur. La publication elle-même peut être la source de l'information, lorsque la cessation et le passage en ligne sont annoncés dans le dernier numéro imprimé.

Profitant de cette situation privilégiée, GDP recueille et analyse les données relatives aux cessations en faveur d'une publication en ligne depuis 2007. La saisie des données dans un tableau<sup>48</sup> permet d'analyser les tendances de la migration des périodiques en ligne. Chaque mois, on indique le nombre de publications migrées ainsi que le pourcentage par rapport au nombre de cessations enregistrées durant ce mois, le mode de diffusion (en ligne ou par courriel), la classe Dewey à laquelle se rapporte la publication, ainsi que des informations concernant la publication imprimée d'origine, notamment le fichier de bulletinage (fichier général ou fichier des PO) et le département de conservation destinataire. Toutefois, ces statistiques sont encore récentes et loin d'être exhaustives, puisqu'elles ne concernent bien entendu que les publications migrées que le service a été en mesure d'identifier. Il convient de les analyser avec précaution, et surtout de pérenniser l'expérimentation afin d'affiner cette analyse, qui entre en correspondance avec l'objectif du DDL, inscrit dans le contrat de performance 2009-2011, d'étudier l'édition française et ses mutations.

En 2009, le service a recensé 249 titres migrés (soit 10% des cessations enregistrées), contre 106 en 2007 (5,3% des cessations) et 166 en 2008 (7,6% des cessations), soit une nette progression. Ce sont essentiellement les revues de sciences sociales qui passent en ligne, à hauteur de 51% du total des migrations en 2009 (contre 68% en 2007 et 79% en 2008). Par exemple, le bimestriel *La vie des idées* de édité par la République des idées paraît uniquement en ligne après le numéro 24 de juillet-août 2007. On constate une assez nette érosion de la part des sciences sociales entre 2007 et 2009, même si elles restent fortement majoritaires. Celle-ci se fait au profit des revues de sciences naturelles, mathématiques et sciences appliquées, à hauteur de 40,2% en 2009 alors qu'elle n'était que de 24% en 2007 et 11,40% en 2008. On peut

<sup>49</sup> http://www.laviedesidees.fr, consulté le 01/09/2010



DEVIDAL Maryline | FIBE T | Projet professionnel personnel | 2010

<sup>&</sup>lt;sup>48</sup> Voir en Annexe 2 le tableau statistique des migrations enregistrées à GDP

citer par exemple la revue *Images des mathématiques* <sup>50</sup> éditée par le CNRS qui passe totalement en ligne en 2008, ou encore la série des *Bulletins climatologiques mensuels* <sup>51</sup> édités par Météo France. Parallèlement la proportion de documents diffusés par courriel était de 16% en 2007, 6,6% en 2008 et 4,8% en 2009, soit un net recul au profit du modèle de diffusion en ligne. Il s'agit essentiellement de lettres d'actualités, comme par exemple *La lettre du Syndicat national de l'édition*. <sup>52</sup>

Ce que l'on note surtout, c'est la forte proportion des publications officielles, toutes classes confondues. Une PO est un document édité par un service public, aussi bien un service central qu'un service déconcentré, territorial ou un établissement public. Pour des commodités de bulletinage, en raison des changements fréquents de titre ou des appellations, les PO éditées par les administrations territoriales ont été isolées dans un fichier distinct du fichier général. Les autres PO sont bulletinées avec les autres périodiques dans le fichier général. En 2009, 22% des titres migrés appartiennent au fichier des PO (contre 27% en 2007 et 29% en 2008, soit une légère baisse), alors même que le fichier des PO ne représente que 5% du fichier général dans Millennium. Les périodiques édités par les collectivités publiques semblent donc migrer davantage que les périodiques classiques, surtout si l'on considère que ce chiffre ne concerne que les PO territoriales.

Un deuxième tableau recensant les titres ayant migré vers le web<sup>53</sup> nous donne des indications sur leur périodicité. Le nombre de titres recensés dans ce tableau est plus important que le nombre de titres sur lequel a porté l'étude précédente, puisqu'il a été régulièrement alimenté depuis le début de l'année 2010. Sur 631 lignes de titres 54, 162 titres ont une périodicité mensuelle, soit 25% du total (dont environ 78 éditions départementales du Bulletin climatologique mensuel<sup>55</sup> publié par Météo France, ainsi qu'une vingtaine de Recueils des actes administratifs<sup>56</sup> édités par les préfectures de région). La seconde périodicité en termes de nombre de titres est la périodicité annuelle puisqu'elle concerne 133 titres, soit 21% du total (dont une vingtaine de Kompass région<sup>57</sup> ainsi que de nombreux rapports annuels ou d'activités, budgets ou statistiques émanant d'organismes publics). Parmi les titres restant, 86 ont une périodicité trimestrielle, 42 bimestrielle et 32 paraissent trois fois par an, soit encore 25% du total. Les bimensuels, hebdomadaires et quotidiens restent très minoritaires avec respectivement 16, 9 et 5 titres, et se partagent le quart restant (par exemple le bimensuel Bulletin officiel du Ministère de l'écologie et du développement<sup>58</sup> ou le quotidien la Revue quotidienne de presse<sup>59</sup> éditée par le Comité des constructeurs français d'automobiles). Il semble à première vue que les publications qui migrent vers le web sont majoritairement des titres à périodicité longue ou au moins mensuelle, les périodicités supérieures à un mois représentant 71% du total des titres enregistrés, ce qui s'explique en partie par le poids important des PO.

<sup>&</sup>lt;sup>50</sup> http://images.math.cnrs.fr, consulté le 01/09/2010

http://climat.meteofrance.com/chgt\_climat2/climat\_france, consulté le 01/09/2010

http://adherent.sne.fr/pages/inscription-newsletter.html, consulté le 01/09/2010

Voir en Annexe 3 un extrait du tableau par titres des publications migrées tenu par GDP

<sup>&</sup>lt;sup>54</sup> En date du 1<sup>er</sup> juin 2010

<sup>&</sup>lt;sup>55</sup> http://climat.meteofrance.com/chgt\_climat2/climat\_france, consulté le 01/09/2010

<sup>&</sup>lt;sup>56</sup> Disponibles sur les sites des différentes préfectures

<sup>&</sup>lt;sup>57</sup> <u>http://fr.kompass.com</u>, consulté le 01/09/2010

http://www.bulletin-officiel.developpement-durable.gouv.fr, consulté le 01/09/2010

http://www.ccfa.fr/revue-de-presse/france, consulté le 01/09/2010

Ce tableau nous apprend également, à la simple lecture des titres qui y sont recensés, que la presse d'association et la presse professionnelle représentent une part importante des titres migrés. Il n'est pas possible de donner une estimation car les filtres proposés ne sont pas pertinents pour les isoler.

### Enjeux pour le dépôt légal des périodiques

Les enjeux en termes de gestion des collections de périodiques du dépôt légal, qu'ils soient imprimés ou numériques, sont l'exhaustivité et la continuité des collections, la conservation pérenne des documents, ainsi que la garantie de l'accès aux documents après leur disparition des circuits de distribution. Les publications qui migrent vers le web ou qui naissent en ligne échappent à la collecte menée par le DDL, qui n'est alors plus en mesure de constituer une collection patrimoniale complète. C'est pourquoi le département est actuellement en quête de solutions, lui permettant de mettre en œuvre la collecte des périodiques en ligne.



# UN NOUVEAU CHAMP D'APPLICATION POUR LE DEPARTEMENT DU DEPOT LEGAL

### Définition et cadre juridique

Le champ d'application du dépôt légal des imprimés, institué par François ler en 1537, n'a cessé depuis lors d'être étendu aux nouveaux supports de l'information: estampes, photographie, musique imprimée, audiovisuel, multimédia, etc. Depuis le 1<sup>er</sup> août 2006<sup>60</sup>, la BnF a pour mission de collecter, conserver et communiquer les sites web du domaine français<sup>61</sup>. La responsabilité en a été confiée au Service du dépôt légal numérique (DLN), intégré au DDL. Ce service compte 5 agents qui travaillent en étroite collaboration avec 4 ingénieurs et experts fonctionnels au sein du Département des systèmes d'information. Le DLN est membre du Consortium international pour la préservation de l'internet<sup>62</sup>, et participe à des activités de coopération et de valorisation scientifique.

Deux textes fondent le dépôt légal de l'internet et introduisent les exceptions au droit d'auteur permettant sa mise en œuvre : la loi 2006-961 du 1<sup>er</sup> août 2006 relative au droit d'auteur et aux droits voisins dans la société de l'information (DADVSI), Titre IV relatif au dépôt légal<sup>63</sup>, et le Code du patrimoine, Livre 1<sup>er</sup>, Dispositions communes à l'ensemble du patrimoine culturel, Titre III relatif au dépôt légal<sup>64</sup>.

Le champ d'application du dépôt légal de l'internet est défini comme suit à l'article L131-2 du Code du patrimoine suit à l'égal les signes, signaux, écrits, images, sons ou messages de toute nature faisant l'objet d'une communication au public par voie électronique ». Il s'agit d'une conception élargie du web, incluant notamment les publications diffusées par courrier électronique. Seuls en sont exclus la correspondance privée et les espaces privés hébergés sur le web, c'est-à-dire les intranets. Cela signifie également que ce dépôt ne repose pas sur des critères de contenu ou de qualité, cette absence de discrimination et de choix préalable amenant la BnF à collecter l'ensemble de la production française.

L'exception au droit d'auteur introduite par la loi DADVSI octroie à la BnF, en tant que dépositaire du dépôt légal, le droit de reproduire une œuvre sur tout

http://www.legifrance.gouv.fr/affichCodeArticle.do;jsessionid=F02CF5508C4BD4E1C733CD7CB751748C.tpdjo05v\_2?cidTexte=LEGITEXT000006074236&idArticle=LEGIARTI000006845516&dateTexte=20100902&categorieLien=cid#LEGIARTI000006845516, consulté le 01/09/2010



 $<sup>^{60}</sup>$  Loi relative au droit d'auteur et aux droits voisins dans la société de l'information

<sup>&</sup>lt;sup>61</sup> A l'exception des sites relevant du domaine de la communication audiovisuelle dont la collecte est confiée à l'Institut national de l'audiovisuel

<sup>62</sup> http://netpreserve.org, consulté le 01/09/2010

http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000266350, consulté le 01/09/2010

<sup>64</sup> http://www.legifrance.gouv.fr/affichCode.do?idArticle=LEGIARTI000006845515&idSectionTA=LEGISCTA000006159 934&cidTexte=LEGITEXT000006074236&dateTexte=20090917, consulté le 01/09/2010

support et par tout procédé technique, en vue de sa collecte, de sa conservation et de sa consultation sur place, sans avoir à recourir à l'autorisation explicite de l'auteur de chaque œuvre. C'est cette exception qui rend possible les actes de reproduction effectués au moment de la collecte du web, et la conservation des sites ainsi copiés. Le droit de la propriété intellectuelle est par ailleurs garanti par la limitation de la consultation des archives du web, sur place, et par des « chercheurs dûment accrédités » 66.

En attente du décret d'application de la loi, actuellement en cours de validation, la BnF adopte une conception élargie du territoire national applicable à l'internet, et définit comme d'origine française les sites enregistrés sous le nom de domaine .fr. ou tout autre nom de domaine enregistré auprès des organismes français chargés de la gestion de ces noms, et/ou produits sur le territoire français ou enregistrés par une personne domiciliée en France. Un accord conclu avec l'Association française pour le nommage internet en coopération (AFNIC)<sup>67</sup> permet à la BnF de disposer de la liste exhaustive des noms de domaine en .fr et .re (France et île de la Réunion).

Toutefois, l'internet est un espace documentaire qui ne se laisse pas aisément archiver. Son évolution très rapide et le volume exponentiel de sa production le rendent potentiellement illimité. L'objectif d'exhaustivité, qui est celui du dépôt légal sur support, n'y est pas applicable, car il supposerait de prendre tous les documents, à toutes les étapes de leurs mises à jour qui peuvent être considérées comme autant de nouvelles éditions. On peut ajouter que l'unité documentaire n'est plus la même, la granularité se situant, dans le contexte du web, au niveau du site, et non plus du document. Par ailleurs, les outils d'archivage automatique ne peuvent prendre en compte le web invisible<sup>68</sup>. Les archives du web sont nécessairement partielles et lacunaires. L'exigence d'exhaustivité s'est déplacée vers une recherche de représentativité.

La collecte est indissociable de la préservation, au titre de la fonction patrimoniale, et de la communication auprès du public. Traditionnellement, le DDL ne gère pas ces fonctions, puisque les collections imprimées collectées jusqu'à présent sont transmises aux départements de collection, auxquels elles reviennent. Dans le cas des archives du web il n'y a pas d'autre intervenant que le DDL<sup>69</sup>. C'est donc à lui qu'il revient de se poser les questions liées à l'archivage pérenne des données et à l'accès aux documents.

#### Les modalités de collecte des documents

La loi, au lieu d'envisager le dépôt des fichiers par les éditeurs, autorise les institutions mandataires à copier les sites web sans solliciter l'autorisation préalable des éditeurs. Très rapidement le choix de procédures largement automatisées s'est imposé, de façon à ne pas ralentir les robots par la multiplication d'opérations manuelles. On comprend aisément que le dépôt et le traitement unitaire sont incompatibles avec un objectif qui tendrait à approcher à l'exhaustivité. Par ailleurs, le web n'est pas une suite de documents mais un



DEVIDAL Maryline | FIBE T | Projet professionnel personnel | 2010

<sup>66</sup> Notion introduite pour la première fois dans le droit français par la loi DADVSI

<sup>67</sup> http://www.afnic.fr, consulté le 01/09/2010

On appelle web invisible la partie du web accessible en ligne, mais non indexée par des moteurs de recherche classiques (notamment les pages accessibles sur formulaires)
<sup>69</sup> Mis à part le Département des systèmes d'information chargé de la mise en œuvre opérationnelle

tissu organique entremêlé. Ainsi, c'est la conjugaison de ces deux approches qui permet de couvrir l'immensité du web et d'en restituer le caractère dynamique 10, garantissant ainsi l'archivage d'une copie conforme à l'original.

Contrairement au dépôt légal traditionnel, le dépôt légal des sites web n'implique aucune démarche active de la part de l'éditeur. Les collectes sont faites de manière automatique à l'aide du robot moissonneur Heritrix. Le robot fonctionne comme un internaute automatique : il explore le web à partir d'une liste d'URL<sup>71</sup>. et capture les fichiers qu'il découvre en cliquant de lien en lien. La logique du moissonnage se substitue à celle du dépôt. Le projet prend en compte la nature hypermédiatique et dynamique du web et propose un archivage le plus proche possible de ce qui est réellement en ligne. Cette approche, bien que représentant un certain nombre de difficultés techniques et théoriques, est la plus adaptée car c'est celle qui dénature le moins possible le support collecté, contrairement à un archivage des seuls documents ou encore à un archivage statique des pages visitées.

La masse, les limites du web profond, les difficultés techniques et l'extrême volatilité du web empêchent toute collecte exhaustive. La BnF procède donc par échantillonnage, selon des critères visant à assurer la meilleure représentativité possible et selon une approche intégrée qui conjugue trois méthodes complémentaires : les collectes larges automatiques du domaine français, les collectes ciblées résultant d'une sélection humaine et les collectes à procédure négociée.

Les collectes larges ont vocation à couvrir l'ensemble du périmètre du dépôt légal. Annuelles, elles étaient réalisées depuis 2004 en partenariat avec la fondation Internet Archive<sup>72</sup>. Elles ont été totalement internalisées en 2010. Leur caractère non sélectif et massif en fait la méthode d'archivage la plus économique et la plus conforme à l'esprit du dépôt légal. Néanmoins, si l'étendue de cette collecte est très ambitieuse, sa profondeur est limitée par le volume des données à capturer. Le robot dispose d'un budget maximum de 10 000 URL pour chaque site, ce qui n'est pas suffisant pour la collecte de sites à gros gabarit. La collecte reste donc superficielle. Le caractère très ponctuel et superficiel de ces collectes leur a valu le surnom "d'instantanés" du web français.

Les collectes ciblées sont réalisées plusieurs fois par an, à partir d'un nombre limité d'URL ayant fait l'objet d'une sélection selon une politique documentaire 13. Elles sont de deux types : la collecte courante concerne des sites de référence choisis par les bibliothécaires de la BnF dans leur domaine d'expertise scientifique. Les collectes projet sont menées autour d'un événement<sup>74</sup> ou d'un thème <sup>75</sup>. Au-delà de ce critère documentaire, la collecte ciblée doit être réservée à des sites pour lesquels la collecte large n'est pas adaptée, notamment les sites de grande taille ou les sites bénéficiant d'un fort taux de renouvellement incompatible avec une capture annuelle. Le robot dispose cette fois d'un budget de 100 000 URL, il est donc capable de réaliser une collecte beaucoup plus

<sup>&</sup>lt;sup>75</sup> Par exemple, les journaux intimes, le développement durable... DEVIDAL Maryline | FIBE T | Projet professionnel personnel | 2010



<sup>&</sup>lt;sup>70</sup> Hyperliens, contenus multimédias, web 2.0, etc

<sup>71</sup> Uniform ressource locator

http://www.archive.org, consulté le 01/09/201. La BnF a également acquis les archives d'Internet Archive depuis

<sup>20 000</sup> propositions / an

<sup>&</sup>lt;sup>74</sup> Par exemple, les élections de 2002, 2004 ou 2007

performante. Pour chaque URL sélectionnée, on spécifie la profondeur<sup>76</sup> à laquelle doit descendre la capture, ainsi que la périodicité de la collecte<sup>77</sup>, déterminée par la fréquence de mise à jour du site. En 2004 a été constitué, au sein de la BnF, un réseau de correspondants du dépôt légal de l'internet. Il s'agit d'un groupe de bibliothécaires ayant la charge de sélectionner des sites internet représentatifs de leur domaine d'expertise scientifique. En 2010, ce réseau regroupe 85 bibliothécaires. Des partenaires extérieurs peuvent également être associés au travail de sélection des agents de la BnF, notamment les bibliothèques de dépôt légal imprimeur.

Les collectes à procédure négociée concernent des publications que le robot est techniquement incapable de capturer. La BnF met alors en place des procédures de transfert de fichiers spécifiques, en étroite coordination avec le producteur. Une expérimentation a été menée sur l'archivage de la version électronique authentifiée du *Journal officiel de la République française* depuis juin 2005.

Dans le cas d'un abonnement payant la BnF n'est pas supposée devoir acquitter des droits, le dépôt légal étant gratuit. Toutefois, en attente du décret d'application, la BnF n'a pas souhaité collecter des sites dont le contenu était payant<sup>79</sup>.

# Le signalement des données et l'accès aux documents

Le signalement des données répond lui aussi à une logique de masse. Le catalogage des sites copiés est exclu au profit d'une indexation automatique par URL. C'est cette indexation qui permet de naviguer dans les archives comme dans le web «vivant», en cliquant de lien en lien. L'indexation plein texte est un des objectifs de développement les plus attendus par les utilisateurs, qui souhaitent pouvoir rechercher dans les archives comme ils le font dans le web «vivant». Actuellement seules 5% des collections sont indexées en plein texte.

En 2009, les Archives de l'internet contiennent environ 12,5 milliards de fichiers<sup>80</sup>, soit la deuxième collection au monde. Dans l'attente du décret d'application qui fixera de façon formelle les modalités légales de consultation des collections constituées, la BnF expérimente depuis avril 2008 une plateforme d'accès à la base de consultation des Archives de l'internet<sup>81</sup>. La navigation se fait grâce au logiciel libre Wayback machine. L'interface permet de circuler entre les deux espaces : web «mort» et web «vivant». L'« archinaute » a le choix entre plusieurs modes de navigation possible : la recherche par URL, la recherche par mot, et les parcours guidés. Les parcours guidés permettent de mettre en valeur des collections réalisées autour de thèmes particuliers, par exemple : « Le web militant », « S'écrire en ligne : journaux personnels et littéraires », « Cliquer, voter : l'internet électoral ». Conformément aux restrictions imposées par la loi, la consultation se fait sur place uniquement,

<sup>78</sup> http://www.journal-officiel.gouv.fr, consulté le 01/09/2010

http://renet.bnf.fr/jsp/EcranRessourceNum.jsp, consulté le 01/09/2010 (accessible depuis le site de la BnF



<sup>&</sup>lt;sup>76</sup> Page + 1 clic, site, domaine, hôte

<sup>77</sup> Quotidienne à annuelle

<sup>79</sup> Sauf lorsqu'un accord a pu être passé avec le producteur, comme c'est le cas avec Cairn

<sup>80</sup> Les plus anciens remontant à 1996

dans toutes les salles du Rez-de-jardin sur le site François Mitterrand, et sur le site Richelieu.

### Archivage pérenne des collections

Les données collectées sont stockées dans des fichiers ARC, un format container dans lesquels fichiers compressés et métadonnées sont rassemblés. Les fichiers ARC sont ensuite entreposés dans deux réservoirs distincts. L'un est dédié à la consultation et l'autre à la conservation. Les fichiers ARC font l'objet d'une copie de sauvegarde dans le Système de préservation et d'archivage réparti de la BnF (SPAR)<sup>82</sup>. Véritable magasin numérique, sa conception s'appuie sur la norme OAIS (ISO 14721 : 2003), modèle de référence pour un système ouvert d'archivage pérenne d'information. SPAR est bien plus qu'un simple entrepôt de données sécurisé. Il effectue de multiples copies des objets numériques et assure une surveillance continue de l'état des équipements et des médias d'enregistrement des fichiers, permettant d'anticiper les recopies avant la perte définitive. Il permet également, grâce à une reconnaissance précise et complète des formats de données versées, de garantir la continuité d'accès en procédant aux transformations nécessaires en cas d'obsolescence technologique des outils informatiques de restitution. De plus, SPAR permet à tout moment de revenir en arrière pour restituer les objets dans leur format d'origine.

### LA PLACE DES PERIODIQUES EN LIGNE DANS CE DISPOSITIF

En premier lieu s'est posée la question du mode de collecte des périodiques en ligne, c'est-à-dire la question du choix entre le dépôt par l'éditeur d'un fichier numérique, ou la collecte automatique de la publication dans son environnement dynamique. Une expérimentation menée par le DDL en 2006 sur le dépôt légal de fichiers numériques de la presse quotidienne régionale (PQR) ayant abouti à des résultats peu concluants, l'hypothèse d'un dépôt des fichiers numériques a rapidement été écartée, au profit de la collecte automatique dans le cadre du dépôt légal du web. Revenons sur cette expérimentation.

Une dérogation introduite par le décret modificatif 2006-696 du 13 juin 2006<sup>83</sup> prévoit que la BnF peut demander le dépôt d'un fichier numérique se substituant au dépôt du document imprimé, graphique ou photographique. Les documents numériques ainsi collectés doivent être la copie conforme des documents originaux. Ils ont le même statut patrimonial que les documents auxquels ils se substituent, ce qui implique les mêmes contraintes pour la BnF en termes de conservation et de communication. Le DDL a mené une expérimentation de dépôt légal de fichiers numériques de substitution pour la presse quotidienne régionale (PQR), laquelle pose d'importants problèmes de collecte et de conservation en raison de ses éditions multiples et des gros volumes qu'elles représentent, dans l'idée d'une substitution à terme. En effet, depuis 2004, la BnF assure la conservation complète de l'édition principale sur support imprimé

<sup>&</sup>lt;sup>83</sup>http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000816574&dateTexte=#LEGIARTI000006252 109, consulté le 01/09/2010



<sup>82</sup> Voir en Annexe 4 le schéma présentant le fonctionnement de SPAR

de l'ensemble de la PQR, soit 40 titres, et conserve, en plus, l'ensemble des éditions de 20 titres sur microfilms. L'expérimentation a porté sur deux titres représentatifs : Le Populaire du Centre<sup>84</sup> et Ouest-France<sup>85</sup>. Les différentes éditions de ces deux titres ont été collectées et conservées au format pdf. Un certain nombre de problèmes ont été identifiés au cours de cette expérimentation, notamment la question de la conformité du dépôt numérique par rapport à l'édition imprimée réellement diffusée, dans la mesure où des modifications de dernière minute sont susceptibles d'intervenir. Le traitement de ces fichiers numériques pose également des problèmes de contrôle des fichiers et de bulletinage. Des contacts pris avec d'autres titres ont surtout mis à jour d'hétérogénéité procédures des d'ordonnancement des fichiers selon les producteurs, ce qui constitue le principal frein à une généralisation de cette collecte, puisqu'il faudrait alors, pour chaque titre, instruire des procédures particulières de traitement. Par ailleurs, la préservation a également rapidement posé problème. Pour Ouest-France, les livraisons représentaient quotidiennement pas moins de 4 gigaoctets de fichiers pdf haute résolution, lesquels devaient faire l'objet d'une chaîne de préservation spécifique.

D'une façon générale, le dépôt des fichiers numériques ne permet pas de rendre les publications dans leur contexte dynamique, de restituer les liens qui peuvent pointer vers l'intérieur ou l'extérieur, ni les contenus audiovisuels ou de type web 2.0 qui l'enrichissent. La publication n'est que partiellement collectée et n'est pas fidèle à l'original. On peut ajouter également que dans un univers où la fréquence de renouvellement est excessivement rapide, le dépôt des fichiers ne permet pas une réactivité assez forte. Cette alternative semble assez peu adaptée à une collecte à grande échelle, mais elle peut en revanche être utilisée pour faire entrer dans les collections des titres ne pouvant être collectés automatiquement.

D'autre part, collecter le web dans son ensemble par la collecte automatique, tout en réservant un traitement particulier pour les publications en ligne (périodiques, mais aussi e-books, vidéos, etc) pour lesquelles on demanderait un dépôt de la part de l'éditeur, pose des problèmes de cohérence et de lisibilité des procédures, sans compter que la frontière entre un site web et une publication en ligne est extrêmement perméable. Il importe de simplifier au maximum les procédures pour l'éditeur.

Enfin, la mise à disposition des fichiers au public représente une difficulté supplémentaire et nécessiterait d'instruire d'importants et très coûteux développements informatiques, de façon à pouvoir déposer les fichiers dans la base des Archives de l'internet et reconstituer l'architecture permettant d'y naviguer.

Pour toutes ces raisons, le choix a été fait d'envisager la collecte des périodiques en ligne au sein de la collecte globale du web français. Tout périodique diffusé sur un site web relevant du domaine français relève du dépôt légal du web, et est en principe automatiquement collecté lors de la collecte large. Ce type de collecte présente toutefois des inconvénients et n'est pas toujours adaptée lorsqu'on parle de périodiques. Tout d'abord, il s'agit d'une

http://www.ouest-france.fr, consulté le 01/09/2010





-

<sup>84</sup> http://www.lepopulaire.fr/accueil.html, consulté le 01/09/2010

collecte de surface. Les ressources qui nous intéressent peuvent être ancrées profondément au sein du site et être exclues de la collecte large. Par ailleurs, il s'agit d'une collecte annuelle, incompatible avec des publications dont la périodicité est inférieure à un an et ne disposant pas d'une fonction d'autoarchivage.

Le décret d'application du dépôt légal de l'internet est actuellement en cours de validation. L'avant-projet spécifie que « des collectes automatiques complémentaires de certains sites sélectionnés pour garantir la continuité et l'enrichissement des collections du dépôt légal sur support » pourront être effectuées par la BnF. Les périodiques en ligne ont donc tout à fait vocation à être inclus dans ce processus.

# La stratégie actuelle de collecte des périodiques en ligne

Actuellement, les sélections de périodiques en ligne destinés à être collectés dans les Archives de l'internet proviennent de plusieurs sources, sans qu'il existe de réelle stratégie concertée. Il existe néanmoins deux sources spécialisées de sélections de périodiques en ligne : les sélections de GDP, ainsi que les sélections du CIF.

Le CIF dispose d'un panier de sélections dans lequel il signale des titres de périodiques en ligne repérés lors de l'attribution d'un numéro ISSN. En 2010, le CIF a proposé 218 titres en ligne à la collecte. Parmi les priorités établies par le CIF, on peut mentionner le fait que la ressource soit en cours, qu'elle ne soit pas disponible sur support, qu'elle soit de niveau recherche, qu'il n'existe pas de fonction d'autoarchivage et qu'elle soit en accès libre afin d'éviter les difficultés de collecte liées aux formulaires d'authentification. Le CIF ne tient pas de statistiques qualitatives permettant d'analyser les titres proposés.

GDP dispose également d'un panier de sélections <sup>86</sup>. Actuellement, la priorité du service se porte sur le signalement des ressources annonçant une cessation de l'édition imprimée au profit de la diffusion en ligne, et ce de façon à garantir la continuité des collections. Depuis fin 2006, 631 titres migrés ont ainsi été signalés. Un certain nombre de sélections sont également effectuées sur des titres natifs en ligne, mais de façon beaucoup moins importante et non systématique, au hasard des opérations de prospection menées par l'équipe de veille.

Des titres peuvent également être sélectionnés par les bibliothécaires correspondants du dépôt légal du web. Cette sélection est effectuée en fonction de l'intérêt documentaire du site au regard de la spécialité disciplinaire du sélectionneur, et non en fonction de la typologie du contenu.

Les sites recensés dans les Signets de la BnF<sup>87</sup> (dont un certain nombre de périodiques) et dans le Portail des périodiques électroniques A to Z<sup>88</sup> sont archivés lors de la collecte large du domaine français, de même que les portails

http://renet.bnf.fr/jsp/index.jsp, consulté le 01/09/2010 (accessible sur authentification). Contient environ 20 000 titres sur abonnements ou accessibles gratuitement sur le web



-

<sup>&</sup>lt;sup>86</sup> Voir en Annexe 5 un extrait du panier de sélection du DDL (dont les sélections de GDP) validé par le DLN

<sup>&</sup>lt;sup>87</sup> http://signets.bnf.fr, consulté le 01/09/2010.

de revues Cairn<sup>89</sup>, Persée<sup>90</sup> et Revues.org<sup>91</sup>. Les correspondants des départements ont également la possibilité de proposer, au sein de ces portails, des titres à archiver en collecte ciblée. Par ailleurs, le portail Cairn, avec lequel la BnF a établi un accord, fait partie d'un panier « gros sites » qui vise à une collecte la plus complète possible. Au contraire de la collecte ciblée, dans laquelle le robot dispose d'un budget de 100 000 URL par site, en collecte « gros sites » aucun budget limitatif n'est indiqué au robot, lui permettant de copier la totalité du site. Une surveillance est néanmoins effectuée afin d'éviter que le robot ne se trouve pris dans des boucles. Le panier « gros sites » contient d'autres sites intéressant les périodiques électroniques, notamment Légifrance<sup>92</sup>, l'INSEE<sup>93</sup>, et le CNRS<sup>94</sup>.

Une expérimentation est actuellement menée sur la collecte spécialisée des unes électroniques des titres phares de la presse en ligne : presse quotidienne nationale (PQN), presse quotidienne régionale (PQR), et presse magazine, ainsi que des sites tels que Reuters ou des portails d'information. Des groupes de travail « paniers transverses » ont en effet été créés, au niveau de l'établissement, afin d'alimenter la collecte ciblée sur des projets particuliers. Le groupe « actualités » a pour charge de dresser une liste restreinte à une soixantaine d'URL de sites à collecter quotidiennement <sup>95</sup>. Il ne s'agit pas d'une collecte en profondeur mais uniquement d'une collecte de la page d'accueil + 1 clic, soit l'équivalent des unes imprimées.

# Le bilan de ce qui a déjà été collecté dans les Archives de l'internet 96

Un sondage, mené sur les titres répertoriés dans le tableau ayant servi pour les sélections de GDP, permet de se faire une idée quant à la qualité des archives de périodiques en ligne constituées.

L'analyse a été menée sur un échantillon aléatoire de titres. Après exclusion du tableau des publications diffusées par courriel et des titres n'ayant pas été retenus pour la collecte par le DLN, le nombre de lignes passe de 631 à 313. On constate d'emblée que la moitié des titres proposés échappent à la collecte, soit parce qu'ils sont diffusés par courriel, soit parce qu'ils sont inaccessibles au robot en raison de formulaires d'authentification ou de recherche. Au final sur ces 313 titres seuls 184 sont finalement venus alimenter le panier de sélection du DDL validé par le DLN, soit un peu moins d'un sur trois par rapport au nombre total initialement proposé. Il y a donc une énorme perte avant même que la collecte ait démarré.

Un sondage a été effectué dans les Archives de l'internet<sup>97</sup> sur les titres ayant finalement pu être collectés. Le résultat est assez mitigé. D'une façon générale, les titres proposés au format pdf avec cumulatifs des numéros parus sont

<sup>97</sup> http://renet.bnf.fr/jsp/EcranRessourceNum.jsp, consulté le 01/09/2010 (accessible uniquement sur le site de la BnF)
DEVIDAL Maryline | FIBE T | Projet professionnel personnel | 2010 - 32 -



\_

<sup>89</sup> http://www.cairn.info, consulté le 01/09/2010

<sup>90</sup> http://www.persee.fr, consulté le 01/09/2010

http://www.revues.org, consulté le 01/09/2010

<sup>92</sup> http://www.legifrance.gouv.fr, consulté le 01/09/2010

http://www.insee.fr, consulté le 01/09/2010

<sup>94</sup> http://www.cnrs.fr, consulté le 01/09/2010

<sup>95</sup> Voir en Annexe 6 le contenu du « panier actualités »

<sup>&</sup>lt;sup>96</sup> Voir en Annexe 7 un exemple de consultation dans les archives de l'internet (Cairn)

capturés de manière tout à fait satisfaisante dès la collecte large, l'existence d'une fonction d'autoarchivage permettant par ailleurs de faire l'économie d'une collecte ciblée périodique au profit d'une collecte annuelle. En revanche, les publications tirant partie des fonctionnalités du web 2.0 sont plus difficiles à collecter. Par exemple, les contenus multimédias proposés dans l'édition en ligne du Monde 98 sont inaccessibles. Il arrive aussi que si l'environnement a été collecté de façon satisfaisante on ne puisse accéder à la ressource, bien qu'elle ait été l'objet de la sélection initiale. En effet, la politique mise en œuvre est de capter la ressource dans son environnement éditorial et contextuel, ce n'est donc pas son URL qui sert de point de départ à la collecte mais celle du site qui l'héberge, sauf exceptions. Il peut donc arriver que le robot ne pénètre pas assez profondément dans le site et manque la publication, à moins qu'il ne rencontre des problèmes techniques pour l'atteindre. Signalons par ailleurs qu'il est extrêmement difficile de retrouver certaines publications lorsqu'on ne dispose pas de leurs URL précises, en particulier celles hébergées sur des sites de grande taille, au sein desquels elles sont par ailleurs susceptibles de changer d'emplacement. Par exemple, il n'a pas été possible de retrouver le Mensuel du gouvernement<sup>99</sup> ou les *Bulletins climatologiques mensuels*<sup>100</sup> au sein des archives du site du Premier ministre et de Météo France. Parfois la revue n'est que partiellement collectée. Les Cahiers de jurisprudence de l'aide sociale 101, publiés par la Commission centrale d'aide sociale et hébergés sur le site du Ministère du travail, de la solidarité et de la fonction publique, en sont un bon exemple. L'adresse de départ retenue pour ce titre était celle du portail des bulletins officiels sur Légifrance. Ce portail permet de disposer de liens maintenus à jour, la publication pouvant changer d'environnement en fonction des remaniements ministériels. On accède à partir de ce portail à la page de la publication sur le site du ministère, puis à la liste des derniers numéros parus. A partir de cette liste il est possible d'accéder à une bonne partie des numéros, mais pas à tous. Il est possible d'accéder au N°1 de 2007 mais pas au N°6 de la même année alors que les deux numéros sont accessibles depuis la même page et au même format. Signalons que dans ce cas, l'accès aux années via un menu déroulant est correctement restitué dans les archives. Le cas des Recueils des actes administratifs de la Préfecture de Haute-Savoie 102 est également très parlant. Ils ont été convenablement collectés les 25 février et 12 mars 2009, tous les numéros parus de 2002 à 2008 étant accessibles à partir de liens html. En 2010, 33 collectes du site ont été réalisées. Les recueils n'ont pas du tout été archivés à certaines dates, bien qu'à un clic seulement de la page d'accueil du site (collecte du 3 juin par exemple). A d'autres dates, ils ont bien été archivés, mais pour l'année en cours seulement. L'accès aux numéros des années antérieures se fait désormais par l'intermédiaire d'un menu déroulant, qui n'est pas correctement restitué. Dans les archives constituées le 3 mai, les numéros 1 et 2 sont accessibles, tandis que le lien vers le numéro 3 ne fonctionne pas. L'archive est donc plus ou moins satisfaisante selon la date de la collecte, ce qui ne facilite pas la recherche de l'archinaute.

http://www.haute-savoie.pref.gouv.fr/09\_raa/index.php, consulté le 01/09/2010



<sup>98 &</sup>lt;u>http://www.lemonde.fr</u>, consulté le 01/09/2010

http://www.archives.premier-ministre.gouv.fr/villepin/information/mensuel\_gouvernement\_50/index.html, consulté le 01/09/2010

http://climat.meteofrance.com/chgt\_climat2/climat\_france, consulté le 01/09/2010

http://www.travail-solidarite.gouv.fr/espaces,770/famille,774/publications-et-textes-officiels,893/cahiers-deiurisprudence-de-l-aide,6268.html, consulté le 01/09/2010

Un sondage a également été fait sur les portails de revues Persée 103, Revues.org 104 et Cairn 105. Les trois portails sont archivés en collecte large annuelle, certains titres présents au sein de ces portails pouvant faire, par ailleurs, l'objet d'une collecte ciblée. On constate tout d'abord que les formulaires de recherche ne fonctionnent pas. Ce défaut est cependant mineur en termes d'intérêt pour le dépôt légal des périodiques en ligne, l'important restant de pouvoir accéder aux documents eux-mêmes. Ces trois portails proposent un accès par discipline et/ou par liste de titres de revues, correctement restitué dans les archives. Si l'accès à la liste des numéros semble systématique d'après nos expérimentations, l'accès aux sommaires des numéros et aux articles eux-mêmes est assez aléatoire sur Persée et Cairn. indépendamment de la barrière mobile. Il s'agit, dans la terminologie DLN, de « gros sites » pour lesquels la collecte large ne suffit pas en raison de leur gabarit, mais aussi de leur architecture complexe. Afin d'améliorer la qualité de la collecte de ces gros sites, le DLN est amené à travailler de concert avec les producteurs et à passer des accords. C'est le cas actuellement avec Cairn. Les titres ayant pu être consultés de façon satisfaisante ont généralement été ceux ayant fait l'objet d'une collecte ciblée. Revues.org a pu être collecté de façon plus satisfaisante, notamment car il propose une architecture simple.

Le résultat de ce sondage est plutôt mitigé. Il faut néanmoins tempérer tout sentiment de frustration en gardant à l'esprit que le dépôt légal de l'internet n'en est qu'à ses débuts, et que la maturité à laquelle arrive le DLN lui permet d'envisager de nouvelles solutions à plus ou moins long terme.

## UNE REPONSE ADAPTEE MAIS QUI N'EST PAS SANS SOULEVER DE QUESTIONS

L'extension du champ d'application du dépôt légal à l'internet pourrait permettre de mettre en œuvre le dépôt légal des périodiques en ligne, et de répondre aux enjeux principaux de collecte, de conservation et d'accès aux documents, y compris après la mort du titre et sa disparition du web.

Le dépôt légal de l'internet n'a pas été institué dans le but de collecter les publications en ligne, mais le web dans son ensemble. Néanmoins on peut inscrire cette collecte comme un objectif corollaire du DLN. Il convient alors de réfléchir à la dimension que l'on veut donner à cette collecte (continuité des collections, exhaustivité ou échantillonnage), et aux implications qui en découlent (contrôle qualité, signalement, suivi de la vie du titre, accès et valorisation).

Tout ceci nous amène à nous poser les questions suivantes. Compte tenu de leurs spécificités techniques et éditoriales et de leurs liens avec les collections imprimées, les publications en ligne doivent-elles être traitées comme n'importe quel objet du web ou doivent-elles faire l'objet d'une approche spécifique ? Doiton viser l'exhaustivité comme on le fait pour les collections papier, et ce malgré la logique d'échantillonnage visée pour le reste du web? La veille devra-t-elle être menée prioritairement sur les publications migrées dans une logique de

http://www.cairn.info, consulté le 01/09/2010



http://www.persee.fr, consulté le 01/09/2010

http://www.revues.org, consulté le 01/09/2010

continuité des collections imprimées, ou au contraire viser à l'exhaustivité et s'appliquer à tous les types de publications périodiques en ligne ?

# La question de l'exhaustivité de la collecte et de la continuité des collections

Les objectifs d'exhaustivité de la collecte et de continuité des collections semblent difficiles à atteindre, compte tenu de la masse de l'internet, assortie des difficultés techniques que peut rencontrer le robot : pages accessibles sur formulaire d'authentification ou de recherche, profondeur à laquelle se trouve la ressource à collecter dans le site, rapportée au budget d'URL dont dispose le robot, formats et modèles éditoriaux variés, présence de contenus web 2.0, etc.

Un certain nombre de ces difficultés sont liées au fait qu'il n'existe pas de modèle éditorial homogène en ce qui concerne les périodiques en ligne. La fonction d'autoarchivage permettant d'adapter la fréquence de capture n'est pas toujours proposée. Il faut donc impérativement, pour chaque ressource, spécifier des critères de fréquence de capture basés tantôt sur la périodicité du titre, tantôt sur l'existence d'archives, la qualité de celles-ci et la possibilité pour le robot d'y accéder. Certains titres sont gratuits et totalement libres d'accès, donc facilement collectables. D'autres titres sont accessibles uniquement sur abonnement, qu'il soit gratuit ou payant, ce qui les rend inaccessibles au robot en raison des formulaires d'authentification. En principe l'éditeur est tenu de fournir, à la demande de la BnF, les codes ou spécifications techniques permettant au robot d'accéder à la totalité du site. Le DLN n'étant actuellement pas en mesure de réaliser les développements techniques nécessaires, cette solution reste envisageable dans l'avenir. Les formats de diffusion sont également variés, les plus courants étant le pdf et le html.

Par ailleurs, le DLN n'a pas les mêmes visées que le dépôt légal traditionnel. L'exhaustivité et la continuité des collections ne sont pas inscrites dans ses missions. Il vise au contraire à effectuer un simple échantillonnage. Les deux types de collecte, large et ciblée, concourent à cet objectif de façon complémentaire. La collecte large vise à obtenir un cliché superficiel du web français à un instant t. Les collectes ciblées ont un objectif plus fin : établir, au sein de ce cliché de surface, un sondage en profondeur et répété sur un échantillon donné, de façon à obtenir une image plus fine permettant d'appréhender des questions liées à la profondeur et à la durée. La collecte ciblée répond également à une logique liée au contenu. Elle vise à capturer des sites web présentant un intérêt documentaire et scientifique (sélections réalisées par le réseau des correspondants du web), des sites hébergeant des publications présentant un intérêt lié à la continuité des collections (cas des ressources migrées), ou à la recherche de l'exhaustivité (cas des publications natives en ligne et des publications hybrides). Il est donc difficile d'inscrire l'exhaustivité de la collecte des périodiques en ligne parmi les missions du DLN.

Le projet de service 2008-2011 du DDL souligne qu' « il est nécessaire, en s'appuyant sur une veille éditoriale active tant sur l'édition imprimée qu'électronique (...), de poursuivre la politique d'exhaustivité raisonnée de collecte, en reprécisant ses objectifs prioritaires et en la formalisant pour chaque type de documents et à chaque étape : collecte, veille, et signalement ». Inscrire



l'exhaustivité raisonnée, comme objectif de collecte des périodiques en ligne, semble le meilleur moyen en effet de la mettre en œuvre, en fonction des moyens disponibles et des compétences techniques actuelles.

### La question du contrôle qualité

Actuellement, Le DLN n'effectue pas de contrôle qualité sur le contenu des archives. Il effectue simplement des opérations de surveillance en cours de collecte, afin de vérifier que celle-ci n'occasionne pas de gêne pour le producteur d'une part, et afin de vérifier que le robot ne se trouve pas pris dans des boucles d'autre part. Il édite également des rapports statistiques portant sur de grands volumes.

Lorsque nous avons dressé le bilan des archives de périodiques en ligne constituées, nous avons remarqué une perte non négligeable lors de la collecte. Il serait nécessaire de réaliser un contrôle qualité après collecte, afin de vérifier que les ressources visées ont bien été collectées, et qu'elles sont accessibles. Dans le cas inverse, un diagnostic permettrait d'ajuster les paramètres de collecte et d'en améliorer la qualité. Les paramètres ajustables les plus faciles à repérer sont l'URL, le mode de collecte, la profondeur et la fréquence.

L'URL signalée pour la collecte est décisive. Actuellement, la politique est de restituer la ressource dans son environnement. C'est donc l'URL du site hébergeant la ressource qui est donnée au robot, et non l'URL de la ressource elle-même. Ce choix offre par ailleurs la garantie d'une plus grande stabilité, l'URL du site étant moins susceptible de varier que celle de la ressource. Le risque est alors que le robot ne descende pas assez profondément dans le site pour collecter la ressource. Le contrôle qualité effectué pourrait permettre d'ajuster ce paramètre et de proposer les deux URL de façon à garantir la capture de la ressource.

Un mauvais diagnostic de la profondeur ou de la fréquence de la publication peut conduire à un mauvais paramétrage et à un mode de collecte non adapté. Le contrôle qualité permettrait de réajuster ces paramètres, notamment de proposer un titre en collecte ciblée plutôt que large.

La collecte peut également échouer pour des raisons techniques, notamment la rencontre par le robot de formulaires ou mots de passe non détectés au moment de la sélection. Les titres une fois identifiés pourraient faire l'objet de négociations portant sur la reconnaissance de l'adresse IP de la BnF, permettant ainsi au robot d'accéder à ces documents.

# La question de l'archivage pérenne

Comme toute archive du web les périodiques en ligne collectés ont vocation à rejoindre l'entrepôt numérique SPAR pour y être archivés de façon pérenne. La question de l'archivage pérenne ne soulève pas de difficulté particulière.

# La question du signalement

La question du signalement des périodiques en ligne se décline suivant deux problématiques, d'une part la mise à jour de la notice de l'édition imprimée pour



les titres ayant migré en ligne, d'autre part la rédaction d'une notice propre à l'édition en ligne.

La mise à jour des notices décrivant des périodiques imprimés ayant migré en ligne est indispensable. Il s'agit en réalité de cataloguer la mort du périodique imprimé et de mettre à jour les produits bibliographiques. La notice est close, et une note publique annonçant la cessation et la migration est ajoutée. Cette note pourrait être enrichie par la saisie d'une URL cliquable permettant d'accéder au document sur le web «vivant». Cependant cela ne va pas sans poser de questions. En effet, l'URL est une information particulièrement instable, la ressource pouvant changer d'emplacement dans le site, le site pouvant changer d'adresse, la ressource pouvant changer d'éditeur, etc... L'expérience permet de constater que l'URL des revues savantes est en général assez stable, le producteur mettant en œuvre une vraie politique éditoriale incluant la pérennité de l'accès à la ressource. En revanche, les URL des revues de collectivités publiques, d'associations, etc. sont beaucoup moins stables, en particulier parce qu'elles sont liées à la vie de leurs producteurs. C'est toutefois le choix de l'URL précise de la ressource, bien que la plus volatile, qui nous semble le plus pertinent dans le contexte de la notice bibliographique décrivant l'édition imprimée. Il s'agit en fait, à ce stade, de saisir une note historique signalant la migration vers le web, la date à laquelle elle a eu lieu et la première occurrence URL constatée. En revanche, s'il existe une fonction d'autoarchivage des anciens numéros, c'est l'adresse de la page de sommaire qui sera privilégiée. L'adresse du site hébergeur pourra être donnée lorsqu'il n'existe pas de page de sommaire mais seulement un lien vers le numéro en cours sur la page d'accueil. Il convient de souligner à ce stade que les URL ne sont pas cliquables à ce jour dans le Catalogue général de la BnF. Ce développement est inscrit au planning des futurs développements du catalogue, notamment dans le cadre du proiet de refonte du catalogue, qui vise à élargir le catalogue général aux autres catalogues et bases descriptives de la BnF et à multiplier les relations entre ces bases, à ouvrir le catalogue sur l'extérieur par l'interopérabilité et le référencement, et à exposer ses données dans le web sémantique.

La rédaction d'une notice propre à l'édition en ligne est en revanche facultative, et soulève la question de la pertinence de signaler des collections que l'on ne possède pas. Par ailleurs, la masse supposée de documents à traiter conduit à envisager un tel signalement avec prudence. Le catalogue général de la BnF contient un certain nombre de références de documents non conservés dans les collections de l'établissement. En ce qui concerne plus particulièrement les périodiques, le CIF, à titre d'exemple, attribue régulièrement à la demande des éditeurs des numéros ISSN à des titres imprimés non déposés, ainsi qu'à des titres en ligne, et rédige les notices bibliographiques correspondantes. On peut argumenter par ailleurs, que signaler des documents en ligne qui sont collectés dans les Archives de l'internet revient effectivement à signaler des collections que possède la bibliothèque, l'intérêt étant pour l'usager d'en avoir connaissance et de pouvoir y accéder.

A partir du moment où l'on décide de cataloguer les périodiques en ligne se pose la question du suivi de la vie du titre et de la mise à jour des informations bibliographiques, dont l'URL. Comme tout périodique, le périodique en ligne est susceptible de changer de titre, de cesser de paraître temporairement ou définitivement, de changer d'éditeur, d'adresse, etc. Il semble là encore

impossible d'assurer une veille signalétique sur tous les titres paraissant en ligne, lesquels au contraire des titres imprimés ne sont pas bulletinés donc examinés par un opérateur humain susceptible de repérer ces changements. Il existe des outils capables de vérifier automatiquement la validité de listes d'URL, toutefois ils ne sont pas complètement fiables. Par exemple, une indisponibilité temporaire du serveur hébergeant la ressource conduit à un diagnostic d'URL invalide. Inversement les rapports d'erreur sont considérés comme des URL valides. Par ailleurs, l'URL peut être toujours valide mais la publication être défunte. Même si la performance de ces outils est relative, on peut estimer qu'elle est suffisante. De tels outils sont par ailleurs utilisés par les outils de gestion d'abonnements électroniques, tels que le portail A to Z, et le CIF envisage d'en acquérir un pour réaliser le suivi des URL saisies dans les notices correspondant aux ressources en ligne ayant reçu un ISSN.

# La question de l'accès et de la valorisation

L'accès au document se fait dans un premier temps dans le web « vivant ». La saisie d'une URL de migration dans la notice de l'édition imprimée, ainsi que le catalogage de l'édition électronique, sont intéressants en cela qu'elles permettraient de diriger le lecteur vers la publication dès la consultation du catalogue. Une valorisation des ressources en accès libre serait également envisageable via le Portail des périodiques électroniques A to Z<sup>106</sup>, moyennant une sélection préalable conforme à la politique documentaire des périodiques électroniques. Le portail A to Z, édité par Ebsco, dispose d'un outil de gestion de contenu incluant la description des ressources, ainsi que la vérification des URL. Anciennement il existait un lien entre le Catalogue général et le Portail A to Z, par l'intermédiaire de parties d'exemplaires rattachées aux notices des périodiques dans le catalogue. Il aurait été très intéressant d'utiliser ces parties d'exemplaires pour signaler les titres ayant fait l'objet d'une valorisation dans le portail. Cependant, il a été décidé tout récemment de les supprimer car elles étaient très lourdes à gérer. Une valorisation est également envisageable dans les Signets de la BnF<sup>107</sup>, là encore conformément à la politique documentaire des signets. Seuls les titres libres d'accès et faisant l'objet d'un site propre y seront éligibles, les signets ne proposant pas de lien vers le contenu ou partie de sites mais vers les sites eux-mêmes. A l'intérieur des différentes disciplines, il existe une rubrique « Revues » permettant de lister des titres de périodiques en ligne. Le contrôle des URL est effectué par les correspondants signets ayant signalé les ressources.

A partir du moment où la ressource n'est plus disponible sur le web « vivant », la question est de savoir comment rediriger le lecteur vers les Archives de l'internet, qui deviennent alors le seul moyen d'accéder à la ressource. Le lien entre le Catalogue général de la BnF<sup>108</sup> et la plate-forme d'accès aux Archives de l'internet<sup>109</sup> pourrait se faire par l'intermédiaire d'une partie d'exemplaire saisie dans la notice bibliographique, ce qui rend l'hypothèse du catalogage des périodiques en ligne vraiment intéressante et interdit de l'exclure a priori. Une

http://renet.bnf.fr/jsp/EcranRessourceNum.jsp, consulté le 01/09/2010 (accessible uniquement sur le site de la RnF)



<sup>106</sup> http://renet.bnf.fr/jsp/index.jsp, consulté le 01/09/2010 (accessible sur authentification)

http://signets.bnf.fr, consulté le 01/09/2010

<sup>108</sup> http://catalogue.bnf.fr, consulté le 01/09/2010

note d'avancement sur le dépôt légal de l'internet, datée du 22 juillet 2009 et adressée à la direction générale de la BnF, souligne en effet que « s'il n'est pas concevable de cataloguer les archives à la pièce, des liens pourraient être imaginés avec le catalogue pour certains types de sites ». La saisie d'une telle partie d'exemplaire devra se faire après contrôle de la collecte, c'est-à-dire après s'être assuré que l'archive existe bien. Les Archives de l'internet ne sont consultables que sur place et offrir un lien inactif à partir du catalogue public pourrait se révéler frustrant pour les usagers. Toutefois, il est à noter que comme tout dépôt légal, celui de l'internet est assorti d'une restriction de la consultation sur place, et que de la même façon, l'usager peut consulter au catalogue des notices de périodiques imprimés, auxquels il ne peut accéder de chez lui.

Le projet de refonte du catalogue permettrait, à partir d'une seule requête, d'avoir une liste de tous les points d'accès à la publication : dans le Catalogue général de la BnF, dans le web «vivant», dans le portail A to Z et les signets, et enfin dans les Archives de l'internet.

La valorisation des périodiques en ligne au sein des Archives de l'internet n'est pas chose aisée. Actuellement, les archives se présentent comme un ensemble hétéroclite de documents dans lequel il est difficile de naviguer. Compte tenu de la masse, le catalogage est exclu au profit d'une indexation par URL, seule une minorité de sites bénéficiant aujourd'hui d'une indexation plein texte. Cela ne va pas sans soulever de difficultés en ce qui concerne les périodiques en ligne collectés, car il faut disposer de l'URL exacte de la ressource, correspondant à la date de la collecte, pour pouvoir espérer y accéder. La mise à jour de l'URL dans la notice du catalogue est donc malheureusement aussi cruciale qu'utopiste. Le DLN a inscrit le développement de l'indexation parmi ses priorités pour 2011-2012, et participe activement aux développements informatiques internationaux, dans les domaines de l'indexation plein texte et de la recherche intelligente. Toutefois, il convient de souligner que l'indexation des périodiques au sein des archives, par titre et par ISSN, ne serait quère pertinente. Une requête lancée sur un titre obtiendrait autant de réponse qu'il existe d'archives pour ce titre, soit énormément de bruit. Il n'est pas possible non plus actuellement au sein des archives de faire des sélections par filtres pour isoler un type de document. Créer un parcours du web « Périodiques en ligne » n'est pas non plus envisageable car ce type de parcours est figé dans le temps.

# Enjeux déclinés par origine de la publication

Les enjeux sont différents suivant que la publication est native en ligne, migrée ou hybride.

### Titres natifs numériques

Diffusées sans recours à aucun support physique et sans lien avec des collections imprimées, les publications nées numériques échappent totalement à la collecte traditionnelle menée par les services du DDL. L'objectif idéal serait donc de parvenir à collecter ces publications de manière aussi exhaustive que



possible, et d'en assurer le suivi. En réalité les choses sont loin d'être aussi simples.

Une des nombreuses questions soulevées par cette étude est justement de cibler les enjeux liés à ce type de publications, et de délimiter le ressort de responsabilité au sein de l'établissement. En principe, ces publications nées numériques et sans lien avec les collections imprimées sont des objets du web à part entière, et leur collecte relève du DLN. Les enjeux ne sont alors plus du tout les mêmes que pour les périodiques imprimés, le DLN n'ayant aucune mission d'exhaustivité, de continuité, et de signalement des collections. Toutefois, on peut choisir de considérer que les périodiques en ligne, mêmes nés numériques, ne sont pas des objets standards et nécessitent un traitement à part. Ce traitement serait alors à envisager comme une activité transverse à l'échelle de l'établissement, coordonnée par GDP. L'objectif de l'exhaustivité de la collecte serait à nouveau de mise, à condition de le rationaliser et de tendre à une exhaustivité raisonnée.

La question du signalement se pose en termes de pertinence, de faisabilité et de moyens. Il serait extrêmement intéressant de pouvoir décrire ces ressources dans le catalogue général et de proposer à l'usager des liens cliquables vers la ressource dans son contexte web et vers les Archives de l'internet, mais concrètement cela ne semble pas possible en raison de la masse supposée de documents à traiter. Cependant, le CIF étant amené à attribuer des ISSN à des ressources en ligne, une partie de ces publications pourraient faire l'objet d'un catalogage. La politique d'attribution d'ISSN aux périodiques en ligne du CIF, jusque là très prudente, repose sur des critères restrictifs en fonction de l'intérêt documentaire, ainsi que de la valeur scientifique de la publication, et l'ISSN n'est attribué que sur demande de l'éditeur. Le CIF souhaite aujourd'hui orienter sa stratégie vers une démarche plus ouverte et plus prospective. L'idée serait de mettre en place une cellule de veille et d'aller au-devant des éditeurs en leur proposant l'attribution. Les périodiques en ligne ayant fait l'objet d'un catalogage par le CIF pourraient faire l'objet d'une valorisation dans le portail des périodiques électroniques A to Z et dans les signets de la BnF, la sélection étant déjà effectuée en amont.

#### Migrations en ligne

Les périodiques ayant migré de l'imprimé vers le web soulèvent la question de la continuité des collections. La collecte des titres migrés semble donc devoir être aussi exhaustive que possible afin de compléter la tête de collection imprimée.

En termes de signalement, il est a minima indispensable de clore la notice bibliographique de l'édition imprimée et de signaler le passage en ligne, aux fins de mise à jour du catalogue et des produits bibliographiques. L'URL correspondant à la date de la migration, réelle ou constatée, pourra avantageusement être mentionnée.

La rédaction d'une notice correspondant à l'édition en ligne et liée à la précédente permettrait d'enrichir le catalogue de façon intéressante. Le nombre relativement restreint de migrations enregistrées depuis la fin de l'année 2006 (631) permet d'envisager de telles opérations, de même qu'une valorisation dans le portail A to Z et dans les signets, ainsi qu'un contrôle qualité portant sur des échantillons représentatifs. Non inscrit dans les missions de la BFP, le

catalogage de l'édition en ligne serait à la charge du CIF, qui a par ailleurs inscrit l'attribution d'ISSN aux publications migrées dans la liste de ses priorités. La veille sur la vie des publications en ligne, et donc la mise à jour des notices, est extrêmement difficile à mettre en œuvre, les fascicules n'étant pas bulletinés. Pour le moment le nombre de titres migrés recensés reste restreint mais il va continuer de croître et les couches de notices à vérifier vont rapidement se superposer. On peut cependant envisager de réaliser un contrôle a minima des URL afin de garantir la validité des liens saisis.

En sus des enjeux habituels se pose le problème de la mise à jour de l'application de bulletinage, les cardex devant impérativement être clos afin d'éviter de générer des relances inutiles.

### Ressources hybrides

Dans le cas de ressources hybrides dont le contenu en ligne est identique au contenu papier, l'éditeur est tenu de continuer à déposer la ressource imprimée. En effet, le dépôt légal est lié à la notion de support plutôt qu'à la notion d'œuvre. Ainsi, une même œuvre peut être déposée plusieurs fois, sur différents supports. L'enjeu est donc double. Il s'agit d'identifier ces publications dans le but de vérifier la continuité du dépôt de l'édition imprimée, doublé de l'objectif de collecter la ressource en ligne. On peut s'interroger toutefois sur la pertinence de viser à l'exhaustivité de cette collecte, sachant qu'elle vient en doublon. Il semble plus opportun en effet de la réserver à un certain type de publications.

Il existe un certain nombre de catégories de périodiques imprimés pour lesquelles le traitement et la conservation sont particulièrement lourds à gérer, et qui sont disponibles parallèlement en ligne. Lorsque les publications sont disponibles à l'identique en ligne, il paraît intéressant de chercher à remplacer le dépôt légal de l'édition imprimée par celui de l'édition numérique, la solution consistant pour le DDL à considérer le dépôt légal du web comme un dépôt numérique de substitution. C'est la réflexion qui est actuellement menée au sujet des PO (par exemple les Recueil des actes administratifs des préfectures), de la PQR, des journaux d'annonce gratuits déclinés en x éditions régionales (par exemple Paru vendu<sup>110</sup>), des publications à feuillets mobiles, ainsi que des normes AFNOR<sup>111</sup>. Il convient dans ce cas de mettre à jour la notice de l'édition imprimée afin de mentionner la substitution de la collecte et la date à laquelle elle intervient, ainsi que le cardex dans Millennium. Le catalogage de l'édition en ligne et l'ajout d'une partie d'exemplaire vers les Archives du web sont également à envisager.

Il existe un autre type de ressources hybrides. Il arrive que le contenu en ligne soit différent du contenu de l'édition imprimée. L'édition en ligne peut proposer des contenus rédactionnels différents ou englober des contenus audiovisuels ou de type web 2.0, comme c'est le cas pour Le Monde 112. Elle peut également être complémentaire de l'édition imprimée. Il s'agit donc d'œuvres distinctes, pour lesquelles la pertinence d'une double collecte est réelle. Cependant, le nombre de ces publications est difficilement évaluable. Il est surtout extrêmement



<sup>&</sup>lt;sup>110</sup> <u>http://www.paruvendu.fr</u>, consulté le 01/09/2010

http://www.boutique.afnor.org/NEL1AccueilNormeEnLigne.aspx?&nivCtx=LIVZLIVZ1A10A201&ts=9081393, consulté le 01/09/2010 112 http://www.lemonde.fr, consulté le 01/09/2010

difficile de les repérer car il faudrait être capable, dans un premier temps de repérer toutes les publications hybrides, puis d'identifier celles dont le contenu est complémentaire à l'édition imprimée en comparant les deux éditions pour chaque titre. Une veille et une collecte spécifique de ces publications ne semble par conséquent pas envisageable.

# Lettres d'information électroniques diffusées par courriel et titres disponibles sur abonnement

Les documents diffusés par courriel sont des documents particulièrement difficiles à collecter, en raison des formulaires d'identification d'une part, mais surtout en raison de leur mode de diffusion. Le DLN a mené une expérimentation sur ce type de documents. Une boîte mail générique a été créée afin de s'abonner à un certain nombre de titres. Ce mode de collecte pose un certain nombre de problèmes : modalités d'abonnement, hétérogénéité du contenu des courriels 113, difficultés d'identification des titres, d'archivage et d'accès aux documents. Cette expérimentation a finalement été arrêtée, principalement en raison des problèmes d'accès. En effet, les documents ainsi collectés et stockés dans la boîte mail sont totalement inexploitables et incommunicables. L'autre hypothèse qui consiste à demander le dépôt du fichier numérique à l'éditeur, n'est guère plus satisfaisante, puisque là encore l'accès à la publication n'est pas possible en l'absence de plateforme dédiée.

Les compétences technologiques actuelles du robot ne lui permettent pas de franchir la barrière des formulaires d'authentification et donc de collecter les titres disponibles sur abonnement, qu'il soit gratuit ou payant. Ces titres doivent être signalés à part afin de faire l'objet de négociations portant sur la reconnaissance de l'adresse IP de la BnF, permettant ainsi au robot d'accéder à ces documents.

#### Bilan

Deux logiques se dégagent : celle de la continuité des collections et celle de l'exhaustivité de la collecte. On peut choisir de privilégier la continuité par rapport à la tête de collection imprimée, ce qui implique une veille exhaustive sur les publications migrées et rend envisageable un traitement signalétique complet. En termes de continuité des collections hybrides, l'édition numérique pourrait être avantageusement collectée en lieu et place de l'édition imprimée, lorsque la collecte de l'édition imprimée pose problème. On peut au contraire choisir de privilégier l'exhaustivité de la collecte, c'est-à-dire élargir la veille et la collecte aux publications nées numériques ou hybrides. Ce scénario impliquerait de délimiter des champs prioritaires afin de viser à une « exhaustivité raisonnée » et de pratiquer un signalement allégé.

Les moyens dont dispose le département pour assurer cette nouvelle mission, impliquent d'envisager des solutions pesant le moins lourd possible sur les équipes, qui doivent continuer de collecter et traiter les documents imprimés, dont le nombre continue de croître. Les stratégies mises en place devront autant que possible être réalisables à moyens constants.

<sup>&</sup>lt;sup>113</sup> Par exemple, les courriels ne contiennent pas systématiquement la publication, il s'agit parfois d'un lien html vers le site web



11

# Eléments de comparaison

# PROBLEMATIQUES RENCONTREES PAR D'AUTRES SERVICES DE LA BNF ATTRIBUTAIRES DU DEPOT **LEGAL: LES LIVRES ET LES DOCUMENTS AUDIOVISUELS**

# Le Service de gestion des livres (GDL)

L'avènement du livre numérique est régulièrement annoncé depuis la fin des années 1990, avec la commercialisation des plates-formes et des tablettes de lecture. Les succès commerciaux mitigés connus par ces produits ont depuis démontré qu'un phénomène de bascule rapide et totale du monde de l'édition vers le numérique n'était guère envisageable à court terme. Toutefois, l'entrée dans ce secteur des géants Amazon, Sony et Google, et les récentes innovations technologiques apportées sur les tablettes de nouvelle génération (Kindle, IPad) relance le débat. La BnF s'interroge en retour sur la meilleure stratégie à adopter pour intégrer la collecte du livre numérique dans la filière du dépôt légal. Le livre numérique s'entend ici comme l'objet numérique, et non comme le support de lecture, tel que défini dans le Rapport Patino 114 sur le livre numérique. Le rapport Patino englobe beaucoup d'objets sous le vocable de « livre numérique », des ouvrages proposés par Numilog 115 aux rapports pdf mis en ligne, en passant par les livres de Gallica<sup>116</sup>. Toutefois, la définition la plus fréquemment retenue reflète une forte analogie avec le livre imprimé contemporain, et implique la présence d'une page de titre, d'une pagination continue, d'un sommaire ou d'une table des matières. On tend également à différencier « livre numérique » et « livre numérisé ».

Certains types de contenus semblent plus concernés par la bascule numérique : l'édition professionnelle, les dictionnaires et encyclopédies, les livres pratiques, etc. D'autres sont susceptibles de connaître à leur tour un phénomène de migration : les livres scolaires 117 et les ouvrages universitaires. La littérature générale, ou fiction, est elle aussi de plus en plus concernée, notamment grâce à la numérisation de fonds d'éditeurs comme Gallimard. Pour autant, ce domaine semble davantage exploré par les diffuseurs sous la forme d'un modèle de couplage imprimé/numérique que d'une véritable bascule. La chronologie de diffusion ne se fait plus exclusivement de l'imprimé vers le numérique, mais on voit désormais des succès nés numériques faire l'objet, dans un second temps, d'une publication imprimée. Le secteur de l'autoédition connaît également un essor important sur internet : des éditeurs spécialisés proposent aux auteurs intéressés de vendre leurs ouvrages en limitant le plus possible les coûts de

Expérimentation en cours sur le « cartable numérique »



<sup>&</sup>lt;sup>114</sup> PATINO, Bruno, Rapport sur le livre numérique : rapport au Ministre de la culture et de la communication, 30 juin 2008, consultable en ligne à l'adresse http://www.culture.gouv.fr/culture/actualites/conferen/albanel/rapportpatino.pdf (consulté le 01/09/2010)

http://www.numilog.com/accueil.asp, consulté le 01/09/2010

http://gallica.bnf.fr, consulté le 01/09/2010

tirage. Soit l'ouvrage est imprimé à la demande, soit il est commercialisé directement en ligne sous la forme d'un e-book à télécharger.

Une grande part de l'édition en ligne concerne donc des œuvres déjà entrées dans les collections patrimoniales sous forme imprimée. Compte tenu des ressources disponibles, la BnF se tourne plutôt vers la collecte de contenus inédits. A ce titre, l'attention se porterait prioritairement vers les plates-formes d'autoédition. L'idée serait de substituer le dépôt imprimé par une collecte des sites distributeurs.

Par ailleurs, le dépôt légal des e-books doit viser les usages qui y sont associés ainsi que le contexte dans lequel ils sont proposés, c'est-à-dire les fichiers, la plateforme, ainsi que l'ensemble des fonctionnalités proposées.

Une expérimentation de collecte a été menée, en mars 2009, sur les modalités décrites ci-dessus. Cette expérimentation s'est limitée aux plates-formes accessibles gratuitement, le DLN ne souhaitant pas prendre le risque juridique de la capture de contenus payants en l'absence du décret d'application sur le dépôt légal de l'internet. Le principal écueil constaté lors de cette expérimentation concerne les conditions de communication et de consultation des documents. Les fichiers sont bien collectés mais le formulaire de recherche inclus dans le catalogue de l'éditeur n'est pas opérationnel dans l'univers des Archives de l'internet et les archives elles-mêmes ne sont pas indexées. Il est impossible d'effectuer une recherche par titre. Par ailleurs, la variété des formats proposés, si elle n'empêche pas la collecte, est un frein à la consultation car il faut disposer des applications permettant la lecture des documents.

Malgré tout, la collecte des livres numériques par le biais du dépôt légal de l'internet constitue une alternative économique qui permettrait à tout le moins de ne pas manquer les « incunables » du e-book.

# Le Département de l'audiovisuel (DAV)

Le département de l'audiovisuel est particulièrement intéressé à la question du dépôt légal des documents numériques hors support, du fait d'une part de sa responsabilité de gestion des documents numériques sur support, et d'autre part en raison de la migration massive de ses collections vers internet. En revanche, le département ne peut bénéficier du dépôt numérique de substitution qui ne concerne que les documents imprimés.

Les documents numériques considérés présentent des caractéristiques techniques qui complexifient la question de la collecte par rapport aux imprimés : tailles des fichiers, formats, protections éventuelles 118, modes de diffusion 119...

En ce qui concerne le son, il n'y a pas jusqu'à présent de baisse du nombre de dépôts qui serait corrélée avec la migration de support. Néanmoins, les tendances actuelles sont préoccupantes. La priorité pour le département est donnée aux plateformes de téléchargement dans l'optique de la continuité des collections pour l'édition commerciale. De nombreux sites proposent des documents sonores en ligne à télécharger : Fnac<sup>120</sup>, ITunes<sup>121</sup>, Virgin<sup>122</sup>, sites

http://www.fnac.com/telecharger-musique.asp, consulté le 01/09/2010



11

<sup>&</sup>lt;sup>118</sup> Les Digital Rights Management

<sup>119</sup> Essentiellement streaming ou téléchargement

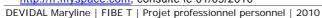
de maisons de disques... Jusqu'à présent, les documents en ligne restent majoritairement diffusés en parallèle des documents sur support, néanmoins la tendance à la substitution s'accélère très rapidement, ainsi que la part de documents inédits sur support. L'autoproduction se développe de manière très massive sur internet via par exemple des plateformes de streaming vidéo telles que Daily Motion<sup>123</sup> ou des sites d'espaces personnels comme. My Space<sup>124</sup>.

En ce qui concerne l'image animée, on n'observe pas de baisse du nombre de dépôts sur support. Néanmoins, plus en détail, la situation est contrastée entre une baisse réelle des nouveaux titres édités par les majors, une hausse sensible du nombre de titres de DVD de l'édition commerciale indépendante, et le passage rapide de la communication institutionnelle vers la diffusion en ligne. Pour ce qui est des plateformes de vidéo à la demande. l'offre reste largement parallèle à l'offre sur support. La vidéo à la demande commence néanmoins à émerger comme moyen de diffusion d'œuvres n'ayant pas accès à l'édition sur supports. La question de la conservation pérenne de ces documents se pose en cas de collecte, du fait des protections DRM dont elles peuvent faire l'objet. Une collecte des fichiers avant apposition des DRM serait nécessaire. Les vidéos institutionnelles représentent environ 40 % des dépôts sur supports. On assiste pour ce type de document à une substitution progressive d'une production sur support à une diffusion hors support. Plusieurs cas de figure coexistent : mise à disposition de fichiers numériques en ligne sur des sites internet accessibles, mise à disposition de fichiers numériques hors de l'espace public de navigation, sur des intranets par exemple. Ces documents semblent donc à la frontière du périmètre du dépôt légal du web, puisque le critère légal d'assujettissement au dépôt des publications électroniques réside précisément dans le fait qu'elles doivent faire « l'objet d'une communication au public en ligne par voie électronique ».

En ce qui concerne le multimédia, les documents électroniques habituellement collectés sur support sont de plus en plus diffusés uniquement en ligne, et accessibles via un téléchargement gratuit ou payant selon les cas. Les documents électroniques sur support sont également parfois à compléter par des contenus en ligne : les bases de données et encyclopédies, les jeux vidéo et les logiciels voient leurs mises à jour diffusées uniquement en ligne, par téléchargement. Ces mises à jour sont indispensables à la complétude du document initial. De manière périphérique à la problématique de la continuité des collections, dans la mesure où il n'existait pas d'édition sur support, les contenus développés pour les terminaux mobiles et accessibles uniquement en ligne par téléchargement constituent également une préoccupation nouvelle.

Le DAV est donc particulièrement impacté par cette question du dépôt légal des documents numériques hors support. La masse de documents concernés et leurs caractéristiques techniques constituent une réelle spécificité et ne trouvent pas pour l'instant de solution de collecte adaptée dans le cadre du dépôt légal du web. Le département souhaite donc que soit instruite la question de la collecte de fichiers numériques.

http://fr.myspace.com, consulté le 01/09/2010







http://www.apple.com/fr/itunes/download/, consulté le 01/09/2010

http://www.virginmega.fr/musique.htm?estat\_svc=s%3d295095185287%26class%3donglets%26page%3donglet%2bmusique%26type%3d40%26g%3d229029185286, consulté le 01/09/2010

http://www.dailymotion.com/fr, consulté le 01/09/2010

# Nécessité de mettre en place des stratégies adaptées

Les autres départements recevant le dépôt légal sont confrontés à des questionnements assez similaires à ceux que connaît GDP, les principales problématiques soulevées étant celles de la continuité des collections sur support et de la collecte des œuvres inédites.

Toutefois, en dépit de ces problématiques communes, les difficultés rencontrées liées aux types de documents à collecter sont propres à chaque service, et entraînent des stratégies variées. Le dépôt légal des e-books, bien que relativement simple à mettre en œuvre, est fortement compromis par la quasi-impossibilité d'accéder aux documents au sein des Archives de l'internet. La collecte des contenus audiovisuels en ligne est au contraire extrêmement complexe à mettre en œuvre, compte tenu des difficultés techniques rencontrées. Le DAV se montre plus intéressé par la solution du dépôt de fichiers numériques de substitution, tandis que l'expérimentation d'un tel dépôt menée sur la presse quotidienne ne s'est pas révélée très convaincante. Il ne semble donc pas opportun d'instruire la question du dépôt légal de l'édition en ligne de manière globale, chaque service devant au contraire être libre de déterminer ses propres stratégies.

# APPROCHES ENVISAGEES PAR DES BIBLIOTHEQUES NATIONALES ETRANGERES METTANT EN ŒUVRE DES ACTIONS D'ARCHIVAGE DU WEB

# La Bibliothèque royale du Danemark

La Bibliothèque royale du Danemark s'est vue confier la mission de collecter le web danois le 22 décembre 2004. Elle n'a pas à obtenir l'accord des producteurs pour archiver leurs sites. Elle se trouve donc en charge d'objectifs comparables à ceux de la BnF, et tout comme la Bnf elle a adopté une approche intégratrice combinant trois types de collecte : une collecte large 4 fois par an, une collecte ciblée menée quotidiennement sur 87 sites sélectionnés (dont 43 sites de presse en ligne), et une collecte thématique menée en rapport avec des événements. Le projet se nomme Netarchive.dk<sup>125</sup>.

# La Bibliothèque nationale du Québec

La Bibliothèque nationale du Québec mène une approche extrêmement sélective visant les publications gouvernementales. Cette extrême sélectivité, rendue nécessaire par le fait que la bibliothèque doit obtenir l'accord du producteur du site avant de l'archiver, lui permet de viser à l'exhaustivité dans ce domaine. Depuis 2001, environ 20 000 publications signalées par les ministères et organismes gouvernementaux ont été déposés dans la Banque des publications gouvernementales 126, accessible par internet.

http://services.banq.qc.ca/sdx/pgq/accueil.xsp?db=notice, consulté le 01/09/2010



DEVIDAL Maryline | FIBE T | Projet professionnel personnel | 2010

<sup>125</sup> http://netarchive.dk/index-en.php, consulté le 01/09/2010

# La Bibliothèque nationale suisse

La Bibliothèque nationale suisse a pour mandat légal de collectionner, de répertorier, de conserver et de rendre accessible le domaine suisse de l'internet. La stratégie mise en œuvre est de constituer une collection de « Helvetica » en ligne, de manière progressive et sélective. Pour l'instant, elle collectionne un choix de publications officielles, de publications universitaires, de publications commerciales et de sites web. Pour chaque type de publication, la BN collabore avec d'autres institutions : avec les offices fédéraux pour les publications officielles, avec les bibliothèques des hautes écoles et des universités suisses pour les publications universitaires, avec les éditeurs suisses pour les publications commerciales, et avec les bibliothèques cantonales ainsi que d'autres institutions intéressées pour les sites web patrimoniaux. D'autres domaines de collection viendront s'y ajouter au fil du temps. e-Helvetica 127 est constituée de plusieurs sous-bases : Archives web suisse (sites web), e-Publications (publications officielles), e-Diss (thèses) et POP (publications commerciales). Les sites web sont moissonnés par un robot. En revanche, les publications en ligne sont livrées par leurs producteurs, et accompagnées de leurs métadonnées au format MODS, ce qui suppose un traitement au cas par cas et non une logique d'exhaustivité.

# Absence de projet lié à une typologie bibliothéconomique

On constate que la problématique de la collecte de l'édition en ligne ne s'est pas réellement posée aux autres établissements ayant mis en œuvre des projets d'archivage du web. Nous avons trouvé peu d'exemples de projets incluant un volet « édition en ligne ». Les réflexions menées autour de la collecte du web à l'étranger sont similaires à celle menée actuellement par la BnF : une réflexion autour des typologies de producteurs, de contenu ou de média. La politique menée par la Bibliothèque nationale du Québec sur les PO semble surtout dictée par la volonté de constituer une base de données sur l'information publique, et par la facilité d'obtenir l'accord des producteurs de sites officiels d'autre part. La Bibliothèque royale du Danemark mène une politique d'archivage intégratrice similaire à celle de la BnF, au sein de laquelle une attention particulière est portée à la presse en ligne. Encore une fois cette stratégie particulière semble menée davantage dans le but de collecter l'actualité et son traitement, que les titres de presse eux-mêmes. Le projet de la Bibliothèque nationale suisse est le seul à notre connaissance à adopter une logique bibliothéconomique en réservant un traitement spécifique à l'édition en ligne. Toutefois, cette logique est ensuite déclinée par producteurs (services publics, éditeurs, universités et hautes écoles) et non par typologie de documents. Par ailleurs, le mode de collecte choisi, le dépôt, implique des procédures longues de négociations. Elle offre en revanche la possibilité de constituer des collections de haute qualité, avec la fourniture de métadonnées complètes utiles pour le signalement et la valorisation. Ce projet est tout particulièrement intéressant et pourrait servir de modèle pour des réflexions à venir sur un projet élargi de dépôt légal de l'édition en ligne.



DEVIDAL Maryline | FIBE T | Projet professionnel personnel | 2010



41

La BnF, si elle choisit de se lancer dans l'aventure du dépôt légal des périodiques en ligne, serait donc pionnière dans ce domaine. Libre à elle d'inventer les stratégies, procédures et techniques à mettre en œuvre. Tout d'abord, il semble que mettre en œuvre un tel projet implique d'adopter une vision bibliothéconomique des contenus du web. Les périodiques en ligne ne seraient alors plus envisagés comme n'importe quel objet web, mais comme une publication devant bénéficier d'un traitement particulier. L'approche intégratrice est particulièrement intéressante à ce titre puisqu'elle permettrait de proposer les titres en collecte large ou ciblée (voire en collecte à procédure négociée), suivant la spécificité de chaque publication. Les publications annuelles ou proposant des fonctions d'autoarchivage pourraient être collectées en large, tandis que les publications à périodicité plus rapide et sans fonction d'autoarchivage, les publications ancrées profondément dans les sites ou encore les publications proposées sur abonnement pourraient faire l'objet d'une collecte ciblée ou de négociations portant sur la reconnaissance de l'adresse IP. Cette approche constitue donc le meilleur atout de la BnF pour envisager le dépôt légal des périodiques en ligne.

\*\*\*\*\*\*

Riches des analyses menées plus haut et des comparaisons que nous venons de dresser, nous sommes en mesure de proposer des scenarii pouvant permettre de mettre en œuvre le dépôt légal des périodiques en ligne du domaine français. Nous avons choisi de retenir deux stratégies possibles. Les deux scénarii sont proposés sous forme de fiche pratique, listant les actions à effectuer.

Le premier donne la priorité à l'objectif de continuité des collections, et propose de mener une collecte exhaustive portant sur les seules publications migrées. Cette exhaustivité vise tant la couverture des titres, que la profondeur des archives réalisées. Compte tenu des volumes relativement peu importants, ce premier scénario propose des stratégies de signalement, d'accès et de valorisation des collections enrichies. Ce scénario inclut également la collecte des publications hybrides dont le contenu est identique à celui de l'édition imprimée, et pour lesquelles la substitution de la collecte du document numérique permettrait de résoudre les difficultés liées à la collecte du document imprimé.

L'objectif du second scénario est de mettre en œuvre un réel dépôt légal des périodiques en ligne. Il se fonde pour cela sur une stratégie mettant en avant l'objectif d'exhaustivité de la collecte, tous modèles éditoriaux confondus. En raison des volumes importants, ce scénario propose de viser à une exhaustivité raisonnée et à une collecte de surface. Il propose également des stratégies allégées de signalement, d'accès et de valorisation des collections.

# Scénario A / La problématique de continuité des collections

Problématique : continuité des collections, mise à jour du catalogue et des produits bibliographiques, mise à jour des cardex de bulletinage.

Objectif : exhaustivité de la collecte des publications qui migrent vers le web, ainsi que des publications hybrides pour lesquelles la substitution de la collecte du document numérique permettrait de résoudre les difficultés liées à la collecte du document imprimé.

Responsabilité : GDP pour la veille, les sélections, et la mise à jour de la notice de l'édition imprimée / CIF pour le catalogage de l'édition en ligne / DLN pour la mise en œuvre.

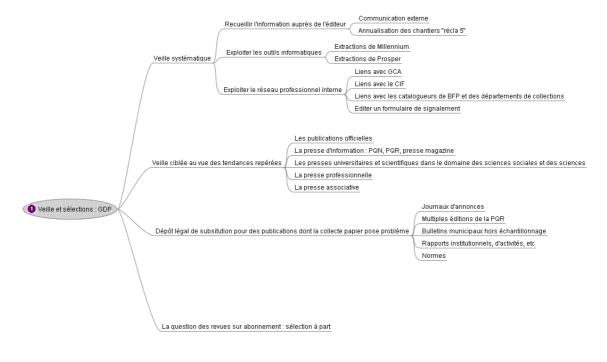
Moyens humains : moyens constants à GDP et au DLN / recrutement d'un ETP supplémentaire au CIF.

# LA VEILLE ET LES SELECTIONS : METTRE EN PLACE UNE STRATEGIE PROSPECTIVE

Réalisées au sein de la cellule de veille de GDP, à moyens constants soit 2 ETP (avec le remplacement du poste vacant).

Actuellement, le recueil d'informations est passif. Les migrations sont repérées fascicules en main lorsqu'elles sont annoncées, à la suite d'un courrier de l'éditeur, ou à la suite de réclamations. Il n'existe pas de veille spécifique sur les publications hybrides.

Mise en place d'une stratégie prospective :





# Veille systématique

## Recueillir l'information auprès de l'éditeur

### Développer la communication externe

Faciliter la communication et l'échange avec l'éditeur: accueil téléphonique, adresse mail générique, courrier postal 28, supports de communication imprimés ou en ligne.

# Aller à la rencontre de l'éditeur : annualisation des chantiers « récla 5 »

Annualiser les chantiers « Récla 5 » et récolter des statistiques de suivi.

(Envisager une période plus favorable que l'été pour ce type de chantiers : taux de non-réponse en partie imputable aux fermetures estivales).

### **Exploiter les outils informatiques**

#### Extractions de Millennium

Réaliser régulièrement des extractions de cardex sur les cases de bulletinage comportant la mention « Cessation. Passage en ligne » 129 => permet de tracer les migrations vers le web.

#### Extractions de l'application de veille Prosper

Exploiter dans Prosper sous Lotus Notus, les réponses enregistrées lors des réclamations de veille<sup>130</sup>. Croiser les champs réponse : « Non » et motif : « En ligne » (le titre paraît de nouveau mais en ligne) => permet de tracer les migrations vers le web.

### Exploiter le réseau professionnel interne

#### Liens avec GCA

Mettre en place une passerelle avec le Service de gestion centralisée des acquisitions : GCA gère les titres entrés par abonnement pour les départements de collections. A ce titre GCA reçoit régulièrement des informations de cessation ou édite des relances : faire suivre à GDP tout ce qui concerne le marché français => permet de tracer les migrations vers le web<sup>131</sup>.

#### Liens avec le CIF

Faire suivre à GDP les informations concernant des migrations repérées lors de l'attribution d'un ISSN à une ressource en ligne.

<sup>131</sup> Ainsi que toute cessation ayant échappé à GDP



<sup>&</sup>lt;sup>128</sup> Voir en Annexe 8 un courrier type relatif aux publications migrées

<sup>&</sup>lt;sup>129</sup> Voir en Annexe 10 le cardex d'un périodique migré dans Millennium

<sup>&</sup>lt;sup>130</sup> Réclamations lancées sur un titre entré dans les collections mais abandonné suite à annonce de cessation, et pour lequel on a une information de résurrection

Créer une passerelle entre les deux cellules de veille (actuellement, le CIF cherche à mettre en place une cellule de veille afin d'orienter sa politique d'attribution vers une démarche plus prospective).

# Liens avec les catalogueurs de BFP et des départements de collections

Notices Val > 1 : au sein des départements de collections des catalogueurs spécialisés dans le traitement des périodiques cataloguent les nouveaux titres entrés par abonnement et mettent à jour les notices : faire suivre à GDP les informations de migration concernant des titres français.

Notices Val = 1 : lorsque la notice est validée au niveau 1 les catalogueurs des départements de collections, non habilités à intervenir à ce niveau de validation, demandent la mise à jour de la notice à la BFP : après mise à jour de la notice, la BFP informe GDP.

### Editer un formulaire de signalement

Sur le modèle des formulaires de signalement des lacunes de titres et de fascicules, mettre en ligne sur Biblionautes<sup>132</sup> un formulaire de signalement de périodiques passés en ligne<sup>133</sup> (dans la rubrique Service public / Infos pratiques / Formulaires communs et dans la rubrique Catalogage / Fiches mémo / Circuits, consignes) => permet de tracer les migrations vers le web.

# Compléter la veille systématique par des chantiers de veille ciblés au vue des tendances repérées

#### Les PO

Mettre en place une veille régulière sur les sites institutionnels nationaux et territoriaux, ainsi que sur les sites des grands établissements publics (221 sites recensés). Veille inscrite dans les missions du profil de poste transverse GDL/GDP « chargé de veille PO ».

Surveiller l'avancée du projet de portail numérique de l'information publique lancé par l'Etat. Effectuer une veille sur ce portail et le proposer en panier « gros sites ».

# La presse d'information : les grands titres de la PQN, de la PQR et de la presse magazine

Veille ciblée sur les éditions électroniques des grands quotidiens conservant parallèlement leur édition imprimée.

Pérenniser l'expérimentation menée sur la collecte spécialisée des unes électroniques de titres phares de la presse en ligne 134, ainsi que des sites tels que Reuters ou des portails d'information (panier « actualités »). Dépouiller

<sup>133</sup> Voir en Annexe 11 un exemple de formulaire possible

La PQN, les titres majeurs de la PQR, la presse magazine
 DEVIDAL Maryline | FIBE T | Projet professionnel personnel | 2010



<sup>132</sup> Intranet de la BnF

notamment Le guide des sites web d'information : les 500 principaux sites d'information français, publié par Développement presse médias en 2010 135.

### Les presses universitaires et scientifiques dans le domaine des sciences sociales et des sciences

Veille régulière sur les sites des grands établissements d'enseignement supérieur, des écoles françaises à l'étranger, des presses universitaires, etc...

### La presse professionnelle

Frange de presse spécialisée, en général magazine, à destination d'une catégorie de professionnels. En général ce type de publications appelle très fortement une diversification des médias : newsletter ou lettre au format pdf pour diffuser l'actualité au plus vite, magazine papier pour la diffusion d'articles d'analyse et bases de données sur le web pour l'accès à des données brutes.

### La presse associative

Remarque : les secteurs prioritaires de la veille sur les publications migrées définis ci-dessus recouvrent en partie ceux fixés pour la veille classique sur les publications imprimées. Il en résulte que les mêmes actions de veille concourent à alimenter les deux filières de collecte, soit un gain de temps et d'efficacité rendant ces procédures tout à fait envisageables dans un avenir proche.

# Mettre en œuvre le dépôt légal de substitution pour les publications dont la collecte papier pose problème

Abandonner la collecte papier de certains types de docs pour lesquels un échantillonnage est suffisant, échantillonage qui pourrait être fait sur les titres en ligne plutôt que papier, ou pour lesquels la collecte de l'édition imprimée pose problème.

Substituer la collecte de l'édition en ligne à celle de l'édition imprimée.

Documents sur lesquels cette substitution pourrait se faire : les journaux d'annonces, les multiples éditions de la presse quotidienne régionale, les bulletins municipaux hors échantillonnage, les rapports institutionnels, les normes, les rapports d'activités, les publications à feuillets mobiles, etc.

# La question des revues disponibles sur abonnement

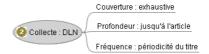
Ne pas exclure de la veille les revues disponibles sur abonnement gratuit ou payant. Réaliser une sélection à part des titres concernés qui pourront faire l'objet de négociations portant sur la reconnaissance de l'adresse IP (attendre le décret d'application pour la capture de contenus payants).

<sup>&</sup>lt;sup>135</sup> Le guide des sites web d'information : les 500 principaux sites d'information français, Paris, Développement presse média, 2010



\_

### MISE EN ŒUVRE DE LA COLLECTE PAR LE DLN



Collecte exhaustive, tant en termes de couverture des titres (veille exhaustive), que de profondeur (jusqu'à l'article) et de fréquence (adaptée à la périodicité du titre, même lorsqu'il existe une fonction d'autoarchivage afin d'éviter les problèmes de menu déroulant non opérationnel dans les Archives du web).

### LE CONTROLE QUALITE DES ARCHIVES CONSTITUEES



Définition des critères de contrôle : succès de la collecte, navigation, contenu (jusqu'au niveau des articles), complétude de la collection (bulletinage).

Contrôle effectué par la cellule de veille de GDP, portant sur des échantillons représentatifs (typologie d'architecture, de périodicité, de formats, etc).

# LE SIGNALEMENT: VERS UN SIGNALEMENT ENRICHI DANS LE CATALOGUE GENERAL DE LA BNF

Actuellement, absence de stratégie concertée entre GDP, la BFP et le CIF:

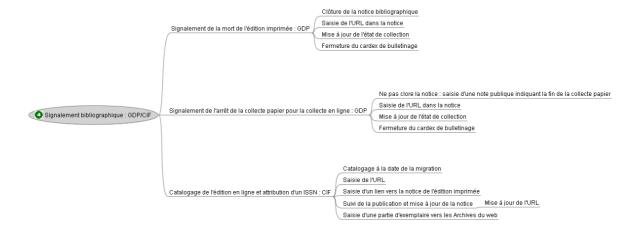
- GDP : fermeture de la notice et du cardex des publications migrant vers le web. Pas de traitement des ressources hybrides.
- BFP : peut être amené à saisir des notes précisant l'existence d'une édition parallèle en ligne lors de la mise à jour de notices d'éditions imprimées (changement de titre, d'éditeur, etc).
- CIF : catalogage complet lors de l'attribution d'un ISSN à une ressource en ligne : mise à jour de la notice de l'édition imprimée et création d'une notice décrivant l'édition en ligne.

L'URL est parfois mentionnée, parfois non. On peut la trouver en zone 300 du format intermarc (visible en affichage public) ou en zone 856 (non visible en affichage public)<sup>136</sup>.

<sup>136</sup> Voir en Annexe 12 des exemples de notices mentionnant une édition en ligne dans le Catalogue général de la BnF DEVIDAL Maryline | FIBE T | Projet professionnel personnel | 2010



### Proposition de stratégie commune et harmonisée :



# Signalement de la mort de l'édition imprimée (cas des publications migrées)

Signaler la continuité de la publication sur un autre média, ainsi que le moyen d'y accéder, et mettre à jour les outils de signalement et de gestion.

Cessations réalisées par GDP à moyen constant : une personne est actuellement chargée de traiter les cessations, y compris les migrations. Il est important que les cessations soient faites par GDP, car il faut également tenir l'application de bulletinage à jour.

Clôture de la notice bibliographique de l'édition imprimée dans le catalogue : saisie d'une valeur codée dans le guide de la notice signalant la mort du titre, d'une date de fin et d'une note publique indiquant la cessation papier au profit de la publication en ligne, ainsi que la date de la migration. Adopter une formulation normalisée en vue de faciliter les recherches éventuelles par mots notice.

Saisie en note publique de l'URL à laquelle est disponible le document (URL à la date de la migration, ou a minima URL à laquelle a été consulté la ressource le jour du catalogage). Dater l'URL.

Mise à jour de l'état de collection dans la partie d'exemplaire de l'édition imprimée.

Clôture de la fiche de bulletinage avec ajout d'une case « Cessation. En ligne », et mise à jour de l'état de collection dans Millennium.



# Signalement de l'arrêt de la collecte papier en faveur de la collecte en ligne (cas des publications hybrides pour lesquelles la collecte papier pose problème)

Tâche supplémentaire à inscrire aux responsabilités dévolues à la personne chargée des cessations à GDP. Réalisable à moyens constants, cette solution de devant concerner qu'un petit nombre de publications a priori.

Ne pas clore la notice de l'édition imprimée dans le catalogue car la publication n'est pas morte. Saisie d'une note publique dans la notice de l'édition imprimée, indiquant le transfert de la collecte vers l'édition en ligne, et la date de la substitution.

Saisie en note publique de l'URL à laquelle est consultable le document. Dater l'URL.

Mise à jour de l'état de collection dans la partie d'exemplaire de l'édition imprimée.

Clôture de la fiche de bulletinage avec ajout d'une case « Cessation. Substitution DL web », et mise à jour de l'état de collection dans Millennium.

# Catalogage de l'édition en ligne et attribution d'un nouvel ISSN

Catalogage réservé aux publications migrées et publications hybrides pour lesquelles on a substitué la collecte de l'édition numérique au dépôt de l'édition papier.

Non inscrit dans les missions de la BFP, le catalogage de l'édition en ligne serait effectué par le CIF. Nécessite le recrutement d'un ETP supplémentaire, pour la reprise du rétrospectif et le suivi et la mise à jour des notices.

Catalogage de l'édition en ligne à la date de la migration ou de la substitution.

Ajout de l'URL à laquelle est consultable la ressource en zone de note publique de façon à ce qu'elle soit visible à l'affichage public (actuellement saisie en zone masquée).

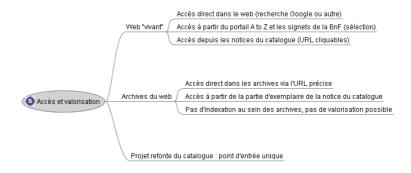
Saisie d'un lien entre la notice de l'édition imprimée et la notice de l'édition en ligne de façon à ce que la notice soit visible à l'affichage public.

Saisie d'une partie d'exemplaire vers les Archives de l'internet (après contrôle qualité, c'est-à-dire après s'être assuré que l'archive a bien été constituée).

Suivi de la publication et mise à jour de la notice de l'édition en ligne à la charge du CIF. Mise à jour minimale : zone de l'URL. Vérifier les URL à l'aide d'un outil de diagnostic de la validité des URL. Afin de rentabiliser l'opération, ce diagnostic sert également à vérifier les URL signalées dans les paniers de sélections validés par le DLN.



# ACCES ET VALORISATION : MULTIPLIER LES POINTS D'ACCES A LA PUBLICATION



# Accès à la publication dans le web « vivant »

La publication est toujours disponible dans le web « vivant ».

Accès direct via une recherche Google ou autre.

Valorisation d'un nombre restreint de publications, choisies en fonction de leur intérêt documentaire et/ou scientifique dans le Portail des périodiques électroniques A to Z<sup>137</sup>, et dans les Signets de la BnF<sup>138</sup> (en conformité avec leurs politiques documentaires). La sélection devra toutefois s'en tenir à un nombre raisonnable de titres, sous peine d'engorger totalement ces outils. L'accès se fait à partir des liens mis à jour.

Accès à partir de l'URL signalée dans la notice de l'édition en ligne et mise à jour (conditionné à l'instruction du développement sur les liens hypertextes dans le catalogue).

# Accès à la publication dans les Archives de l'internet

La publication n'est plus disponible dans le web « vivant ».

Accès directement dans les Archives de l'internet 139, à partir d'une recherche par URL. Pas de valorisation ni de recherche plein texte possibles pour le moment.

Accès à partir de la partie d'exemplaire saisie dans la notice décrivant l'édition en ligne : création d'un lien pérenne pointant vers l'archive constituée à la date de la rédaction de la notice. Possibilité de revenir au calendrier et de choisir une autre date.

http://renet.bnf.fr/jsp/EcranRessourceNum.jsp, consulté le 01/09/2010 (accessible depuis le site de la BnF



<sup>137</sup> http://renet.bnf.fr/jsp/index.jsp, consulté le 01/09/2010

http://signets.bnf.fr, consulté le 01/09/2010

# Projet de refonte du catalogue

Le projet de refonte du catalogue permettra, à terme, à partir d'une seule requête d'avoir accès à tous les points d'accès à la publication : dans le Catalogue général de la BnF, dans le web «vivant», dans le portail A to Z et les signets, ainsi que dans les Archives de l'internet.

#### AVANTAGES ET INCONVENIENTS DU SCENARIO A

Ce scénario, parce qu'il se restreint aux publications migrées et aux publications hybrides pour lesquelles on substitue la collecte en ligne à la collecte imprimée, est facilement réalisable et offre la possibilité d'envisager un traitement complet des documents : collecte, contrôle qualité, catalogage, valorisation et médiation de l'accès. Il satisfait à l'enjeu de continuité des collections puisqu'il permet de ne pas abandonner la collecte de titres jusque là reçus sous leur forme imprimée.

Toutefois, son inconvénient majeur est qu'il ne permet pas de mettre en œuvre un réel dépôt légal des périodiques en ligne car il se restreint justement à un type de publications. Il pose par ailleurs sérieusement la question de la mise à jour des notices, les couches de notices à vérifier étant appelées à se superposer. Par ailleurs, il aboutit à un traitement très discriminant vis-à-vis des publications natives qui en sont totalement exclues, et finalement à une politique peu lisible de l'extérieur. Enfin, il fait peser lourdement la charge du catalogage sur le CIF et implique pour celui-ci de recruter un ETP supplémentaire, alorsmême que le CIF n'a pas de mission de dépôt légal ni de bibliographie.

# Scénario B / La problématique d'exhaustivité de la collecte

Problématique : « exhaustivité raisonnée » à tous les niveaux. En raison des volumes importants de données à collecter il importe d'adopter une stratégie raisonnée, tant au niveau de la veille et des sélections, que de la collecte ellemême (collecte de surface), ainsi que du traitement des documents.

Objectif : collecter les périodiques en ligne de tous types afin d'offrir un panorama de l'édition en ligne de périodiques le plus complet possible, c'est-à-dire mettre en œuvre un réel dépôt légal des périodiques en ligne.

Responsabilité : mission transverse à l'échelle de l'établissement, pilotée par GDP et mise en œuvre par le DLN. Le CIF reste indépendant et maître de sa propre stratégie.

Moyens humains : constants.

# LA VEILLE ET LES SELECTIONS : METTRE EN PLACE UNE STRATEGIE A LA FOIS PROSPECTIVE, CONCERTEE ET RAISONNEE

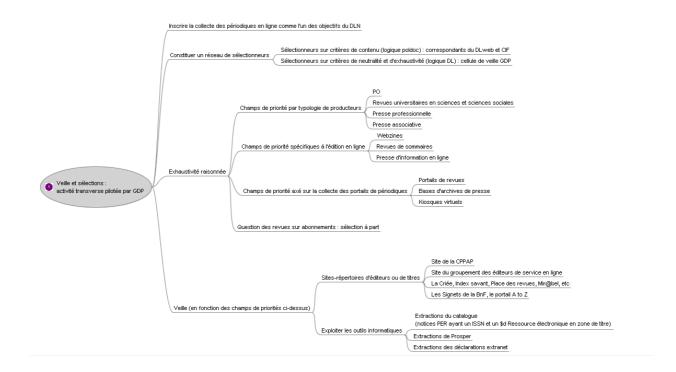
Stratégie actuelle dispersée. Plusieurs origines alimentent les sélections : GDP, le CIF et les correspondants du dépôt légal du web, sans qu'il existe de réelle coordination.

Il appartient à GDP de mettre en place une stratégie de veille et de sélections concertée, garante de la représentativité des collectes réalisées : notamment par l'équilibrage des sélections sur critères par rapport aux sélections neutres, et par la définition de priorités permettant de rendre réalisable une collecte exhaustive.

Les opérations de veille et de sélections sont effectuées à moyens constants car reposent sur des structures déjà en place : cellule de veille de GDP et du CIF, réseau des correspondants du dépôt légal du web.

Remarque : certains éléments de la veille sur les publications migrées décrits au scénario A pourront être repris ici.





# Inscrire la collecte des périodiques comme l'un des objectifs du DLN

### Constituer et animer le réseau des sélectionneurs

Inclure des personnes en charge de sélections sur critères de contenu (les correspondants du dépôt légal du web des départements thématiques et le CIF) et des personnes en charge de sélections sur critères de neutralité et d'exhaustivité (GDP).

# Viser à l'exhaustivité raisonnée en fixant des priorités et en organisant des chantiers ciblés devant alimenter des paniers transverses

Pérenniser et dupliquer l'expérimentation « paniers transverses » menée sur les unes électroniques de la PQN et de la PQR. Plusieurs axes sont possibles pour les futurs paniers transverses en fonction des priorités définies ci-dessous.

# Champs de priorité par typologie de producteurs

Les PO

Les revues universitaires et scientifiques

La presse professionnelle

La presse associative

### Champs de priorité spécifiques à l'édition en ligne

#### Les webzines

Dépouiller les annuaires de webzines tels que Netzine. 140

#### Les revues de sommaires

Publicité scientifique qui fournit gratuitement le sommaire de revues scientifiques payantes publiées sur papier ou en ligne.

Par exemple : *Cultures, Langues, Textes* <sup>141</sup> (revue de sommaires du CNRS), la *Revue de sommaires de la Bibliothèque Cujas* <sup>142</sup>, etc.

La presse d'information en ligne (nationale, régionale, magazine)

# Champs de priorité axés sur la collecte des portails ou bases de périodiques

#### Les bases d'archives de presse

Pressens<sup>143</sup>, Vocale presse<sup>144</sup>, References-indexpresse<sup>145</sup>, etc.

Collectes après négociations portant sur la reconnaissance de l'adresse IP (attendre le décret d'application pour la capture de contenus payants).

### Les kiosques virtuels

Presseregionale.fr<sup>146</sup>, le Portail du Syndicat de la presse régionale, Lekiosque.fr<sup>147</sup>, Monkiosque.fr<sup>148</sup>, etc.

Collectes après négociations portant sur la reconnaissance de l'adresse IP (attendre le décret d'application pour la capture de contenus payants).

DEVIDAL Maryline | FIBE T | Projet professionnel personnel | 2010



-

<sup>140</sup> http://www.netzine.fr, consulté le 01/09/2010

http://www.vjf.cnrs.fr/clt/php/vf, consulté le 01/09/2010

http://webcujas.univ-paris1.fr/principal/rs1/index.php, consulté le 01/09/2010

http://www.pressens.fr, consulté le 01/09/2010

http://fr.vocalepresse.com, consulté le 01/09/2010

http://www.references-indexpresse.com, consulté le 01/09/2010

http://www.presseregionale.fr, consulté le 01/09/2010

http://www.lekiosque.fr, consulté le 01/09/2010 http://www.monkiosque.fr, consulté le 01/09/2010

#### Les portails de revues

Cairn<sup>149</sup>, Revues.org<sup>150</sup> et Persée<sup>151</sup> à inclure dans les paniers « gros sites » (=> bilan et pérennisation de l'expérimentation qui sera prochainement menée sur Cairn).

### La question des revues disponibles sur abonnement

Ne pas exclure de la veille les revues disponibles sur abonnement gratuit ou payant. Envisager une sélection à part des titres concernés qui pourront faire l'objet de négociations portant sur la reconnaissance de l'adresse IP (attendre le décret d'application pour la capture de contenus payants).

#### La veille

Une large partie de la veille s'effectuant sur internet : utiliser les fonctionnalités du web 2.0 tels que les fils RSS afin d'effectuer une veille performante.

Pour le recueil d'information sur les publications migrées, voir scénario A.

### Les sites fournissant des données sur les éditeurs ou les titres

Le site de la CPPAP 152 : depuis la reconnaissance du statut d'éditeur de presse en ligne, les éditeurs peuvent déposer une demande de reconnaissance auprès de la Commission paritaire des publications et agences de presse. Une liste des services de presse ayant été reconnus par la CPPAP est disponible et régulièrement mise à jour sur le site,

Le site du groupement des éditeurs de service en ligne 153 : contient un annuaire des éditeurs en ligne (surtout journaux et magazines imprimés ayant développé une version en ligne),

La Criée 154 : listes de périodiques en ligne dans tous les domaines,

Index savant 155 : répertoire de revues scientifiques en ligne dans le domaine des sciences sociales.

Place des revues 156 : répertoire de revues en ligne,

Mir@bel<sup>157</sup>: réservoir d'informations qui, pour chaque revue recensée, indique où trouver en ligne le texte intégral des articles, les sommaires des numéros, les résumés des articles et les références bibliographiques,

Les Signets de la BnF<sup>158</sup>

http://www.reseau-mirabel.info/index.php, consulté le 01/09/2010



<sup>149</sup> http://www.cairn.info, consulté le 01/09/2010

http://www.revues.org, consulté le 01/09/2010

http://www.persee.fr, consulté le 01/09/2010

http://www.cppap.fr, consulté le 01/09/2010

http://www.geste.fr, consulté le 01/09/2010

http://periodiques.wordpress.com, consulté le 01/09/2010

http://indexsavant.org/index.php?title=Accueil, consulté le 01/09/2010

http://www.placedesrevues.org/index.php/Accueil, consulté le 01/09/2010

# **Exploiter les outils informatiques**

### Extractions du catalogue

Réaliser une extraction rétrospective du catalogue sur les notices contenant un ISSN et un « \$d Ressource électronique » en zone de titre. Filtrages à l'aide du code niveau dans le guide et de la zone de l'indexation Dewey, de façon à obtenir une liste ciblée sur les domaines prioritaires définis, et sur les publications de niveau études et recherche.

### Extractions de Prosper

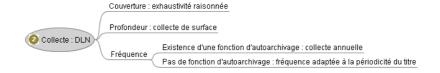
Exploiter dans l'application de veille Prosper les réponses enregistrées lors des prospections lancées sur des publications non reçues au titre du dépôt légal 160. Croiser les champs réponse : « Non » et motif : « En ligne » (le titre existe mais paraît uniquement en ligne) => permet de tracer les publications natives en ligne 161.

#### Extractions des formulaires de déclaration extranet

Exploiter le champ des formulaires de déclaration extranet 162 permettant à l'éditeur de mentionner l'adresse URL de la version en ligne de sa publication dans le cas où elle existe => permet de tracer les publications hybrides.

(Ajouter ce champ aux formulaires de déclaration imprimés).

#### MISE EN ŒUVRE DE LA COLLECTE PAR LE DLN



Collecte dont l'exhaustivité est raisonnée, tant sur la couverture des titres (sélections sur champs de priorité définis), que sur la modalité de la collecte (collecte de surface). La fréquence de la collecte doit être adaptée à la périodicité des titres à chaque fois qu'il n'existe pas de fonction d'autoarchivage. Lorsqu'il existe une fonction d'autoarchivage, la collecte annuelle est préconisée.

<sup>&</sup>lt;sup>162</sup> Voir en Annexe 9 <u>un formulaire de déclaration extranet</u>



<sup>158</sup> http://signets.bnf.fr, consulté le 01/09/2010

http://renet.bnf.fr/jsp/index.jsp, consulté le 01/09/2010

Prospections lancées sur un titre non reçu mais pour lequel on possède des informations attestant de son existence

<sup>&</sup>lt;sup>161</sup> Voir en Annexe 13 une fiche de prospection ayant reçu une réponse « Non : en ligne »

### LE CONTROLE QUALITE DES ARCHIVES CONSTITUEES



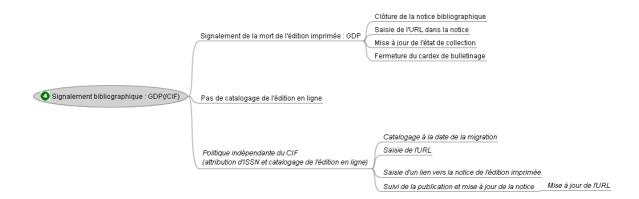
Définition des critères de contrôle : succès de la collecte, navigation, contenu (jusqu'au niveau du sommaire au minimum), complétude de la collection.

Contrôle effectué par la cellule de veille de GDP, portant sur des échantillons représentatifs (typologie d'architecture, de périodicité, de formats, etc).

Prévoir un planning des contrôles échelonné sur l'année, un grand nombre de titres étant destinés à être collectés en large annuelle.

# LE SIGNALEMENT : ADOPTER UNE STRATEGIE DE SIGNALEMENT A MINIMA

Afin que ce scénario soit réalisable et cohérent entre les publications migrées, hybrides et natives, le signalement est envisagé de façon minimale et se restreint à la mise à jour des notices correspondant aux éditions imprimées des publications migrées.



# Signalement de la mort de l'édition imprimée

Mettre à jour les produits bibliographiques et les cardex de Millennium.

Réalisé à GDP, à moyens constants, par la personne chargée des cessations.

Clôture de la notice bibliographique de l'édition imprimée dans le catalogue : saisie d'une valeur codée dans le guide de la notice signalant la mort du titre, d'une date de fin et d'une note publique indiquant la cessation papier au profit de la publication en ligne, ainsi que la date de la migration. Adopter une



formulation normalisée en vue de faciliter les recherches éventuelles par mots notice.

Saisie en note publique de l'URL à laquelle est disponible le document (URL à la date de la migration, ou a minima URL à laquelle a été consulté la ressource le jour du catalogage). Dater l'URL.

Mise à jour de l'état de collection dans la partie d'exemplaire de l'édition imprimée.

Clôture de la fiche de bulletinage avec ajout d'une case « Cessation. En ligne », et mise à jour de l'état de collection dans Millennium.

# Pas de catalogage de l'édition en ligne

# Politique indépendante du CIF : attribution d'ISSN et catalogage de l'édition en ligne

Le CIF reste libre de sa politique d'attribution d'ISSN à des ressources en ligne et de la création des notices associées, telle que mise en œuvre actuellement : à moyens constants.

Ajout de l'URL à laquelle est consultable la ressource en zone de note publique de façon à ce qu'elle soit visible à l'affichage public (actuellement saisie en zone masquée).

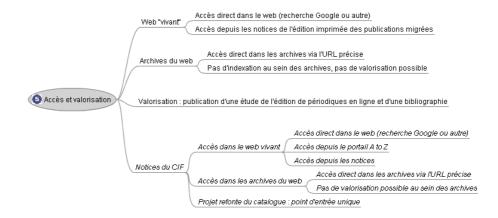
Saisie d'un lien entre la notice de l'édition imprimée et la notice de l'édition en ligne de façon à ce que la notice soit visible à l'affichage public.

Suivi de la publication et mise à jour de la notice de l'édition en ligne à la charge du CIF. Mise à jour minimale : zone de l'URL. Vérifier les URL à l'aide d'un outil de diagnostic de la validité des URL. Afin de rentabiliser l'opération, ce diagnostic sert également à vérifier les URL signalées dans son panier de sélections.

Pas de saisie d'une partie d'exemplaire vers les Archives de l'internet. En effet, cela donnerait une vision trop parcellaire des collections de périodiques en ligne réellement archivées.



# L'ACCES ET LA VALORISATION : ADOPTER UNE STRATEGIE ALLEGEE



# L'accès à la publication dans le web « vivant »

La publication est toujours disponible dans le web « vivant ».

Accès direct via une recherche Google ou autre.

Accès à partir de l'URL signalée dans la notice de l'édition imprimée close, dans le cas des publications migrées. Toutefois, cela ne concerne qu'un nombre restreint de titres et les URL ne sont pas mises à jour.

# L'accès à la publication dans les Archives de l'internet

La publication n'est plus disponible dans le web « vivant ».

Accès directement dans les Archives de l'internet 163, à partir d'une recherche par URL. Pas de valorisation ni de recherche plein texte possibles pour le moment.

Les publications ne faisant pas l'objet d'un catalogage systématique, pas de médiation de l'accès vers la publication via le catalogue, que ce soit dans le web « vivant », ou dans les Archives de l'internet.

# Valorisation auprès de la communauté professionnelle et du public

Tenir des statistiques au moment de la sélection (nombre de titres proposés, modèle éditorial, périodicité, domaine, etc) de façon à recueillir des données quantitatives et qualitatives fines sur les différents types de publications (migrées, hybrides, natives).

http://renet.bnf.fr/jsp/EcranRessourceNum.jsp, consulté le 01/09/2010 (accessible depuis le site de la BnF uniquement)



16

A partir de ces données publier une étude de l'édition de périodiques en ligne dans les pages professionnelles du site de la BnF, rubrique Dépôt légal / Périodiques.

Parallèlement à cette étude, publication d'une bibliographie par le CIF dans les pages professionnelles du site de la BnF (projet à l'étude).

Faire un bilan des opérations de collecte réalisées et proposer un accès aux Archives de l'internet en parallèle de ces deux publications.

# Le cas particulier des titres ayant été catalogués par le CIF

Versement dans le Portail des périodiques A to Z et les signets de la BnF.

Accès à la publication dans le web vivant via le catalogue par l'intermédiaire de l'URL saisie dans la notice de l'édition en ligne et mise à jour (à condition toujours que l'instruction menée sur les liens hypertextes dans le catalogue voit le jour). Accès également depuis le portail et les signets.

Pas d'accès en revanche à la publication dans les Archives de l'internet depuis la notice de l'édition imprimée, puisque la saisie d'une partie d'exemplaire a été exclue.

Le projet de refonte du catalogue permettra, à terme, à partir d'une seule requête, d'avoir accès à tous les points d'accès à la publication : dans le Catalogue général de la BnF, dans le web «vivant», dans le portail A to Z et les signets, ainsi que dans les Archives de l'internet.

#### AVANTAGES ET INCONVENIENTS DU SCENARIO B

Prioriser les sélections, alléger la collecte et renoncer au signalement constituent les principes fondamentaux permettant la réalisation de ce scénario. Celui-ci offre l'avantage, par rapport au scénario A, de mettre en œuvre un réel dépôt légal des périodiques en ligne, et place sur le même plan le traitement réservé aux publications migrées, hybrides ou natives. Le choix de ne pas cataloguer l'édition en ligne écarte par ailleurs la question de la mise à jour des notices.

On peut objecter que le dépôt légal mis en œuvre dans ce scénario relève plus de l'échantillonnage que d'un réel dépôt légal, du moins tel qu'on le conçoit aujourd'hui lorsqu'on parle de collections sur support. Il importe de comprendre que l'échelle du web, sa temporalité et la complexité de sa structure imposent d'adopter une conception nouvelle du dépôt légal. En tout état de cause, il semble plus pertinent pour les chercheurs de demain de disposer d'un panorama élargi de l'édition de périodiques en ligne, même lacunaire en termes de profondeur, plutôt que d'un cliché fin réalisé sur une infime partie de l'offre.

Ce scénario présente toutefois l'inconvénient d'être faible sur les questions de signalement et de médiation de l'accès. Dans l'attente de prochaines évolutions sur les modalités de recherche au sein des archives, les collections constituées sont peu visibles et difficiles d'accès.



# Bilan

	Exhai	ustivité de la ci	ollecte	Conti	nuité des colle	ctions	]				
								Médiation de	Réalisation à		П
								l'accès et	moyens		
	Couverture	Profondeur	Fréquence	Couverture	Profondeur	Fréquence	Signalement	valorisation	constants	Total	
Scénario A	0			2	2	2	2	2	0		10
Scénario B	2	1	1	2	1	1	1	0	2		11

- 2 : objectif atteint 1 : objectif atteint en partie 0 : objectif non atteint



### Conclusion

Les évolutions intervenues dans les modes de diffusion de l'édition de périodiques ne sont pas nouvelles et remontent déjà à une quinzaine d'années pour la presse d'information. Différentes études ont été menées permettant d'en dessiner les contours grossiers. Les données recueillies par les services de la BnF autour des diverses expérimentations de collecte dans le DL du web seront prochainement analysées et feront l'objet d'une étude publiée sur le site. Elles permettront de préciser ces contours et d'ajuster les stratégies de collecte au plus près.

Nous avons dressé un bilan assez mitigé des archives de périodiques en ligne constituées jusqu'à présent. Toutefois, l'expérience et la maturation qu'acquiert le DLN au fur et à mesure qu'il se lance dans de nouvelles expériences lui permettent d'ajuster ses pratiques et d'inventer des solutions techniques allant toujours plus loin. Les évolutions les plus attendues sont celles concernant l'indexation et la valorisation des documents au sein des archives, l'indexation étant cruciale pour la mise en valeur des périodiques collectés, car on l'a vu il est extrêmement difficile actuellement de les retrouver sans disposer d'une URL précise, qui bien souvent correspond à la vie de la publication à un instant t.

Pour autant, il est essentiel de commencer dès à présent à réfléchir à une mise en œuvre du dépôt légal des périodiques en ligne, sans attendre des évolutions techniques qui pourront s'appliquer rétrospectivement aux contenus déjà archivés.

Le scénario A peut sembler le plus urgent à mettre en œuvre car il est garant de la continuité des collections perçues jusqu'à présent sous leur forme imprimée. Il pourrait consister en une première étape du projet, au cours de laquelle seront acquises l'expérience et la maturation nécessaires à la mise en œuvre du scénario B, qui consisterait alors en une montée en charge du projet. En ce cas la question du catalogage de l'édition électronique devra être réexaminée de façon à assurer la continuité de traitement d'un scénario à l'autre.



# Bibliographie

### Législation régissant le dépôt légal

Code du patrimoine, Version consolidée au 14 juillet 2010,

http://www.legifrance.gouv.fr/affichCode.do;jsessionid=762CB5AE8FFD2C95E816ADE C5AEAC267.tpdjo05v\_2?cidTexte=LEGITEXT000006074236&dateTexte=20100819, consulté le 01/09/2010

Code du patrimoine, Livre 1er : Dispositions communes à l'ensemble du patrimoine culturel - Titre III : Dépôt légal,

http://www.legifrance.gouv.fr/affichCode.do?idArticle=LEGIARTI000006845515&idSectionTA=LEGISCTA000006159934&cidTexte=LEGITEXT000006074236&dateTexte=20090917, consulté le 01/09/2010

Décret n°93-1429 du 31 décembre 1993 relatif au dépôt légal,

http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=LEGITEXT000006082758&dateTexte=20090223, consulté le 01/09/2010

Loi n° 2006-961 du 1<sup>er</sup> août 2006 relative au droit d'auteur et aux droits voisins dans la société de l'information,

http://www.legifrance.gouv.fr/html/actualite/actualite\_legislative/decrets\_application/200 6-961.htm, consulté le 01/09/2010

### L'édition numérique

Les chiffres-clés 2008-2009, TarifMédia,

http://www.tarifmedia.com/static/services/infos/chiffres\_cles.pdf, consulté le 01/09/2010 DACOS, Marin, MOUNIER, Pierre, *L'édition électronique*, Paris, La Découverte, 2010, (Repères ; 549)

DACOS, Marin, *Read-write book : le livre inscriptible*, Marseille, Centre pour l'édition électronique ouverte, 2009

JACQUESSON, Alain, RIVIER, Alexis, *Bibliothèques et documents numériques : concepts, composantes, techniques et enjeux*, Paris, Editions du Cercle de la librairie, 2005 (Collection Bibliothèques)

« La presse après les États généraux », Numéro de Regards sur l'actualité, 350, 2009 PATINO, Bruno, Rapport sur le livre numérique : rapport au Ministre de la culture et de la communication, 30 juin 2008,

http://www.culture.gouv.fr/culture/actualites/conferen/albanel/rapportpatino.pdf, consulté le 01/09/2010

POULET, Bernard, *La fin des journaux et l'avenir de l'information*, Paris, Gallimard, 2009 (Le Débat)

« Publications et ressources numériques : les nouvelles pratiques », Numéro de *Archimag. Hors série*, N°35, 2007

SALLES, Daniel, DUFAUT, Olivier, *Du papier à l'internet : les unes des quotidiens*, Futuroscope : SCEREN-CNDP, 2010 (Education aux médias)

TESSIER, Marc, La presse au défi du numérique : rapport au Ministre de la culture et de la communication, février 2007,

http://www.culture.gouv.fr/culture/actualites/rapports/tessier/rapport-fev2007.pdf, consulté le 01/09/2010

TRONQUOY, Philippe, « Information, médias et internet », Numéro de *Cahiers français*, 338, 2007



#### Le dépôt légal du web

GAME, Valérie, ILLIEN, Gildas, « Le dépôt légal d'internet à la Bibliothèque nationale de France : cadre juridique, modèle de collecte, évolutions des métiers », *Bulletin des bibliothèques de France*, 2006, n°3, p. 82-85

ILLIEN, Gildas, L'archivage d'internet, un défi pour les décideurs et les bibliothécaires : scénarios d'organisation et d'évaluation ; l'expérience du consortium IIPC et de la BnF, Communication prononcée à l'occasion du 74ème congrès de la Fédération internationale des associations de bibliothécaires (IFLA), Québec, Canada, 10-14 août

2008, http://archive.ifla.org/IV/ifla74/papers/107-Illien-fr.pdf et

http://www.youtube.com/watch?v=maYtiNcOQEU, consultés le 01/09/2010

ILLIEN, Gildas, « Le dépôt légal de l'internet en pratique : les moissonneurs du web », Bulletin des bibliothèques de France, 2008, n°6, p. 20-27

ILLIEN, Gildas, OURY, Clément, « Quelle politique documentaire pour l'archivage des sites web », dans CARBONNE, Pierre, CAVALIER, François, (dir.), *Les collections électroniques, une nouvelle politique documentaire*, Paris, Editions du Cercle de la librairie, 2009 (Collection Bibliothèques), p. 157-178

#### Sites web

Site de la BnF, rubrique Dépôt légal,

http://www.bnf.fr/fr/professionnels/depot\_legal.html, consulté le 01/09/2010 Site de la DGMIC (Direction générale des médias et des industries culturelles), http://www.ddm.gouv.fr/rubrique.php3?id\_rubrique=24, consulté le 01/09/2010

Site de la CPPAP (Commission paritaire des publications et agence de presse),

http://www.cppap.fr, consulté le 01/09/2010

Site du Consortium IIPC (International internet preservation consortium),

http://netpreserve.org/about/index.php, consulté le 01/09/2010

Site de l'archive danoise Netarchive.dk, http://netarchive.dk/index-en.php, consulté le 01/09/2010

Site de la Banque des publications gouvernementales québécoises,

http://services.banq.qc.ca/sdx/pgq/accueil.xsp?db=notice, consulté le 01/09/2010 Site de l'archive suisse e-Helvetica,

http://www.nb.admin.ch/nb\_professionnel/01693/index.html?lang=fr, consulté le 01/09/2010

#### Documentation interne à la BnF

DELANGE, Olivier, CCDL DL documents numériques, 27 janvier 2009, BNF-ADM-2007-052082-01

ILLIEN, Gildas, Note sur le dépôt légal des « livres numériques » : définition de l'objet et examen des possibilités de collecte automatique, 20 juillet 2009, BNF-ADM-2009-030647-01

JACOBSEN, Hélène, *DDL Projet de service 2008-2011*, 21 juillet 2008, BNF-ADM-2008-026511-01

JACOBSEN, Hélène, *DDL Rapport annuel 2009*, 21 juin 2010, BNF-ADM-2010-012808-01

JACOBSEN, Hélène, *Note d'avancement sur le dépôt légal de l'internet*, 22 juillet 2009, BNF-ADM-2009-041100-01

OURY, Clément, Corr. DLWeb: réunion des coordonnateurs du 14 décembre 2009 (compte-rendu), 5 janvier 2010, BNF-ADM-2009-084965-01

RAISON, Géraldine, DDL réunion sur les publications officielles et le DL web du 2009-09-14, 4 février 2010, BNF-ADM-2009-057368-01

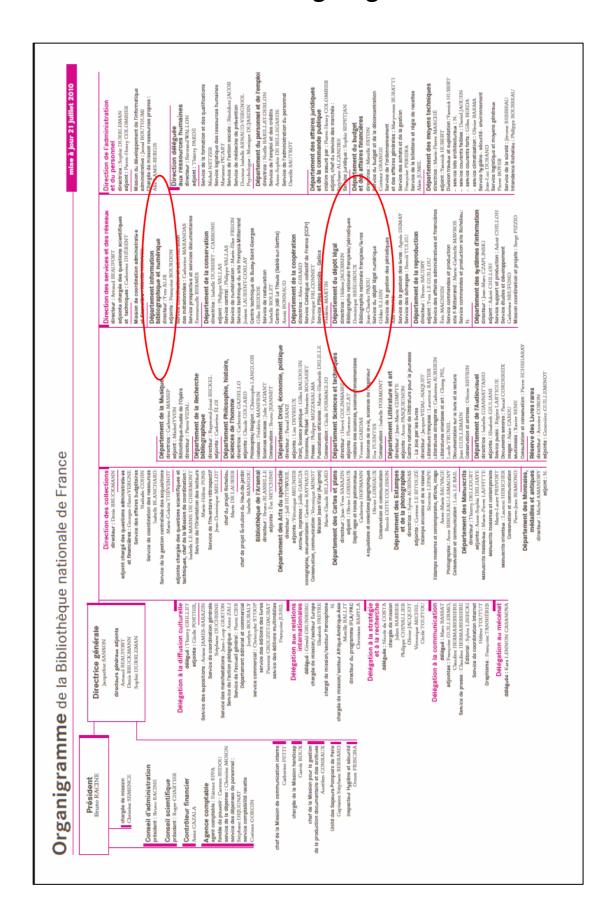


#### Table des annexes

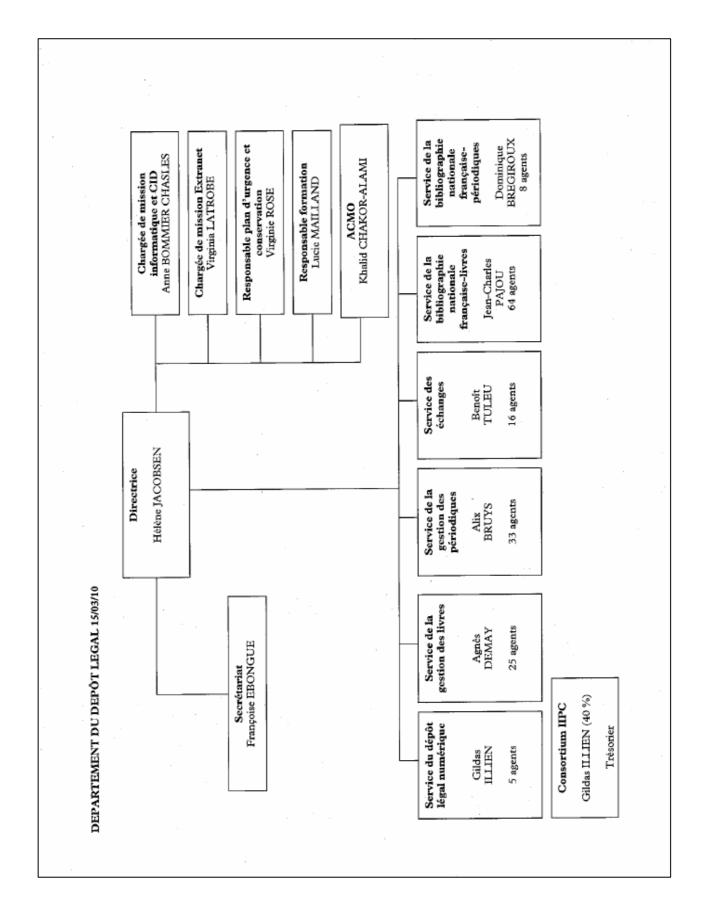
ANNEXE 1 : ORGANIGRAMMES74
ANNEXE 2 : TABLEAU STATISTIQUE DES MIGRATIONS ENREGISTREES A GDP77
ANNEXE 3 : EXTRAIT DU TABLEAU PAR TITRES DES PUBLICATIONS MIGREES TENU PAR GDP83
ANNEXE 4 : SCHEMA PRESENTANT LE FONCTIONNEMENT DE SPAR84
ANNEXE 5 : EXTRAIT DU PANIER DE SELECTION DU DDL VALIDE PAR LE DLN85
ANNEXE 6 : LE CONTENU DU « PANIER ACTUALITES »86
ANNEXE 7 : EXEMPLE D'UNE CONSULTATION DANS LES ARCHIVES DE L'INTERNET (CAIRN)88
ANNEXE 8 : COURRIER TYPE RELATIF AUX PUBLICATIONS MIGREES 95
ANNEXE 9 : FORMULAIRE DE DECLARATION EXTRANET96
ANNEXE 10 : LE CARDEX D'UN PERIODIQUE MIGRE DANS MILLENNIUM 97
ANNEXE 11 : FORMULAIRE DE SIGNALEMENT D'UN PERIODIQUE MIGRE
ANNEXE 12 : EXEMPLES DE NOTICES MENTIONNANT UNE EDITION EN LIGNE DANS LE CATALOGUE GENERAL DE LA BNF99
ANNEXE 13 : UNE FICHE DE PROSPECTION AYANT REÇU UNE REPONSE NON : EN LIGNE104

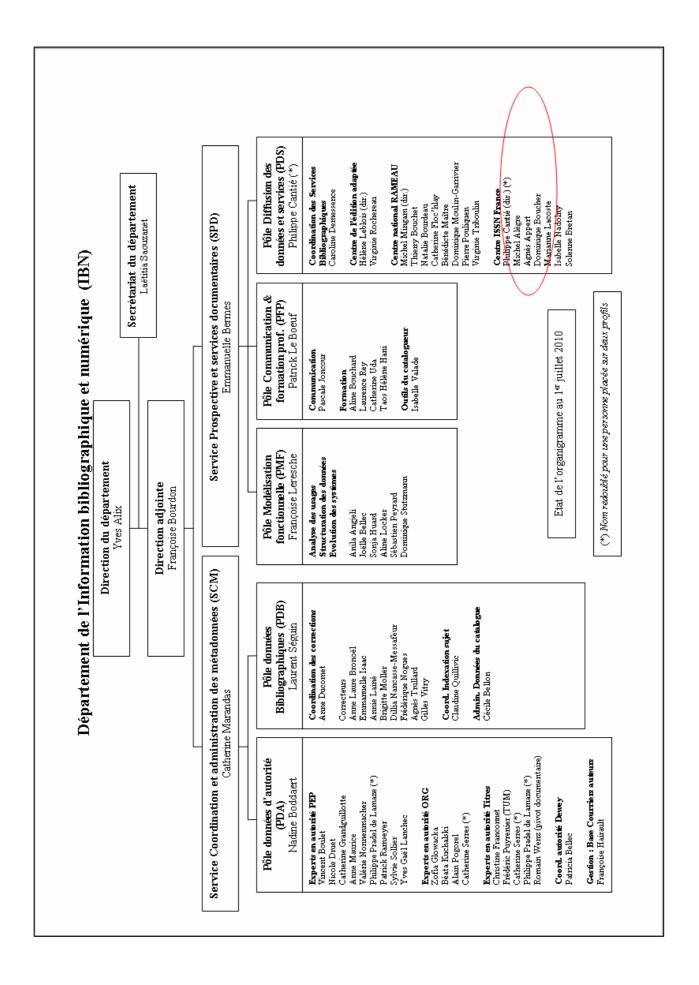


### Annexe 1 : Organigrammes





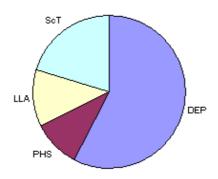




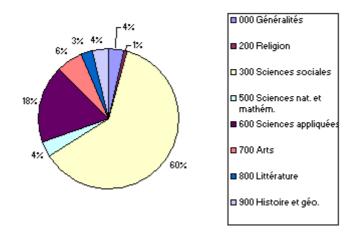
# Annexe 2 : Tableau statistique des migrations enregistrées à GDP

2007													
	janvier	février	mars	avril	mai	juin	juillet	août	septembre	octobre	novembre	décembre	TOTAL
publications migrées	25	4	13	7	8	11	8	7		2	6	19	106
% / cessations de PER imprimés	9,5%	3,6%	8,9%	1,8%	2,8%	4,4%	4,4%	8,9%		0,6%	3,9%	9,7%	5,3%
diffusion en ligne	22	4	10	2	7	9	6	1		2	3	19	81
diffusion par courriel	3	0	3	5	1	2	1			0	3	0	18
fichier FG	11	3	10	7	8	8	5	7		0	6	11	73
fichier PO	14	1	3	0	0	3	2	0		2	0	8	32
% PO	56%	25%	23%	0	0	27%	25%	0		100%	0	42%	27%
DEP	16	2	8	2	4	3	4	6		1	4	12	62
PHS	1	1	3	2	0	2	1	0		0	1	0	11
LLA	1	0	2	2	2	0	2	1		0	0	3	13
ScT	7	1	0	1	2	4	1	0		1	1	4	22
000 Généralités													4
200 Religion													1
300 Sciences sociales													68
500 Sciences nat. et mathém.													4
600 Sciences appliquées													20
700 Arts													7
800 Littérature													3
900 Histoire et géo.													4

Répartition des titres par département de destination en 2007



#### Répartition des titres par classe DEVEY

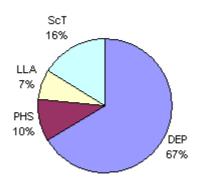




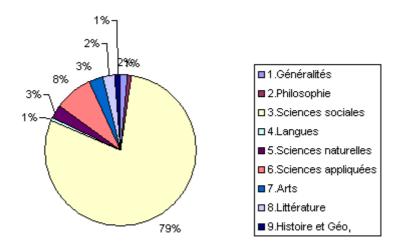
2008													
	janvier	février	mars	avril	mai	juin	juillet	août	septembr	octobre	novembre	décembre	TOTAL
publications migrées	21	18	12	25	8	35	12	4	6	12	10	3	166
cessations PER imprimés	63	279	343	220								1277	2182
% / cessations de PER imprimés	33,3%	6,8%	3,5%	11,0%								7%	7,6%
diffusion en ligne	20	16	10	22	8	33	12	4	6	12	10	2	155
diffusion par courriel	1	2	2	3	0	2	0	0	0	0	0	1	11
fichier FG	12	7	7	18	6	34	5	3	4	5	7	3	108
fichier PO	9	11	5	7	2	1	7	1	2	7	3	0	55
%P0	43%	58%	42%	29%								18%	29%
DEP	16	16	5	17	2	32	8	2	1	6	4	1	110
PHS	1	0	0	3	2	1	2	1	2	3	2	0	17
LLA	0	0	3	1	0	1	1	0	1	2	2	1	12
ScT	4	2	4	4	4	1	1	1	2	1	2	1	27
000 Généralités													3
%													1,#0%
100 Philosophie													1
%													0,60×
200 Religion													•
%													θ×
300 Sciences sociales													131
%													79×
400 Langues													1
%													0,60×
500 Sciences nat. et mathém.													5
%													3×
600 Sciences appliquées													14
%													\$,40×
700 Arts													5
%													3×
800 Littérature													4
%													2,40×
900 Histoire et géo.													2
%													1,20%



## Répartition des titres par département de destination en 2008



#### Répartition des titres par classe Dewey

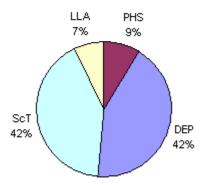


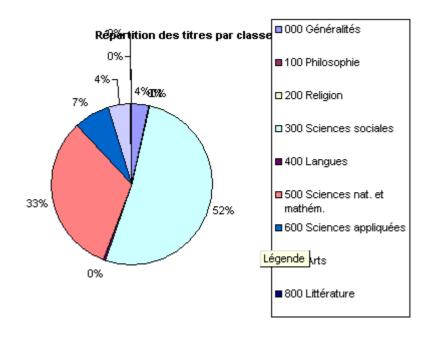


2009													
	janvier	février	mars	avril	mai	juin	juillet	août	septembre	octobre	novembre	décembre	TOTAL
publications migrées	11	4	13	12	19	14	9	2	76	49	30	10	249
cessations PER imprimés	163	107	126	224	202	157	227	110	155	450	266	313	2500
% / cessations de PER imprimés	6,7%	3,7%	10,3%	5,4%	9,4%	8,9%	4,0%	1,8%	49,0%	10,9%	11,3%	3,2%	10,0%
diffusion en ligne	11	3	13	11	17	12	8	2	75	48	28	9	237
diffusion par courriel		1		1	2	2	1		1	1	2	1	12
fichier FG	7	4	11	10	12	6	6	1	76	35	19	7	194
fichier PO	4		2	2	7	8	3	1	0	14	11	3	55
%P0	36%	0%	15%	17%	37%	57%	33%	50%	0%	29%	37×	30%	22%
PHS	1		5		2	1	1	2	1	4	3	2	22
DEP	8	2	5	3	13	12	4		1	35	17	5	105
ScT	1	1	3	7	3		3		73	7	5	1	104
LLA	1	1		2	1	1	1		1	3	5	2	18
000 Généralités	1	1	2						1	2	2		,
%	9,1%												3,6%
100 Philosophie	1												1
%	9,1%												0,4%
200 Religion													
%	0												0%
300 Sciences sociales	7	2	10	3	18	15	5	2	2	37	19	7	127
%	63,6%												51%
400 Langues										1			1
%	0												0,4%
500 Sciences nat, et mathém.				3					73	3	1	1	#1
%	0												33%
600 Sciences appliquées	1	1	1	3			3			4	5		1#
%	9,1%												7,2%
700 Arts	1			3	1		1			1	3	2	11
%	9,1%												4,4%
800 Littérature													
%	0												0,0%
900 Histoire et géo.										1			1
%	0												0,4%



# Répartition des titres migrés par département de destination 2009



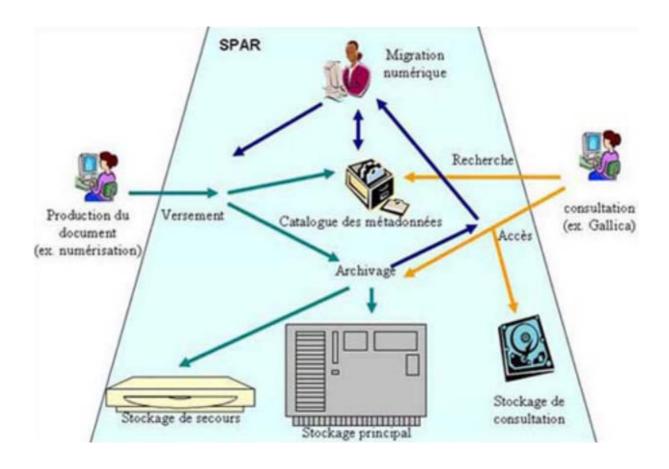




# Annexe 3 : Extrait du tableau par titres des publications migrées tenu par GDP

2hd	Analysis (Marie Marie Ma	PERIODICITE	1	FIN DUFFE	E DVR	DENET	FICHIGN	DEST	Aduly Oper	Well Ables 95 DI COTRONDIE	Security Decides
				H,		Anness stats mig-als					
ZOTZZOWE MAZIG GO SECUTIO	STEP TO THE PROPERTY OF THE PR	20 3	1989	┺		000			en ligne		
2012/2006	U/BILLIATE IMPRESSO de B. U600ce, Service des alcimes	threathe)		208	100 (out, Dec. 2005)	006	2		en Ikna		
and the season of the season o			-			000			en ligne (via baso doc TEME)		
92/01/2007, Si-tex info	1953-6615 Éditions des Parques	trimastrial	1985 X	2006 21	sept. 2006	000	2	a di	900	terral brown	
(22/01/2007) Informatique, bureautique et nouvelles sechnologies	1152-5525 Ministric de la culture	SAINIO	-			009	ē		2	TOTAL CONTINUES AND ASSESSMENT OF THE SECOND	a reveil
TT O400;2007 Bullish official - Mistehn de l'oconomie, des finances, et de l'inst	de 1634-4677 Ministère de l'Aconomie	trimestriei	$\rightarrow$		16	300	2		an ligno	i y neo oenzijesk www	i level
D 0401/2007 Die Pauko	1767-4956 CDOP IContre départemental de documentation pércoccique! d.	L semastriei	1990 20	2008	45 (nov. 2005)	008	ñ	ğ	on ligno	www.crdo.ac-crede#1r	a collection
04/01/2007 Agends du Centre de decumentation	1205-1509 RFI, Centre de documentation	mensuel	1895	11		000	ñ		onligne	www.rlift > actual(de	5
10/01/2007 PASEE lie-de-France falts et childres	1535-9984 NISEE	Simensual	2002	2006	137 (dac, 2008)	300	8		on ligno	High www.insee fallifies replemelatiguabilitains of chilles him	5
OAD1/2007 Avis of publications AFSSA	1769-3179 SMSSSA (Aggree française de cécunité sanitaire)	poune	2001	2003	t 3 (2003)	009	8		en ligno	3,6	3 collector
O401/2007 Memento de statistiques des transports	D758-8719 Ministero des Transports	Sumus	1971 20	2002	2003	300	. 2		en ligne	June inemediate stackment www	, ou
OS91/2007 Bulletin d'informations brèves Asace-Littanie	1770-8907 Assoc-Litanio					300	£	OEP	on ligne	www.fhumie-cuffure blossock.com	nig e
December Lette	1636-6905 L'Officiel	bimensuel	2002	2006	86 (16 déc, 2006)	003	5		countel	info@officialnot.com	
1700/2007 La Letter de 1999	1773-391X Institut français des retations plemationales	Jenzuem	2005	2006	16 (Luin 2006)	906	Ę.	DEP	courriel	Lanlesse@inlorg	6
7) 18/01/2007 Liken	x ESC Rouen		2000	2005 2	14 (kg), 2005)	900	ē	OEP	an figure	ij'Uenna-osa-ednosb',www	à receir
O- 19/01/2006 Dinaba	5771-2106 Ambassade d'Afrique du Sud	memsies	2002	3002		006	5		courrie	maling@afriguesua.net	æ
19/01/2007 Bulletin officiel - Ministère des affaires dagradères	0990.9886 Journaux officiels	Timostrie	1981	2006 25	97 (ect-dec 2005)?	006	2	8	on Kore	the many legitance assume the	d record
19/01/2007 Bulletin official de l'Agence nationale pour l'emplei	1624-8074 Dioction des journaux officies	bimostriel	2000	2006	6 (44c. 2006)	006	2			d) orang series	d const
_	di 1634-9067 Journaux officials	pimensuel		2006	24 (30 ddc 2005)	98	8	ĝ		H TOTAL (TOTAL )	8 1600F
19/01/2007	1766-9022 Ad American	intention	-	↓_	8,046-2008	000	2 8			If also it on the latest the late	noa
_	1286-2848 Direction regionale de Feuicoment le de-France			2000	902	U.S.	. 8				100
T20172027	1298-3497 Cap Wvre ensemble	moresuol		2005 200804	22 (20	500	2	å	o de la co	The state of the s	1908
25/01/2007			-	2005	2005	006	2		atog co	www.arrests.arricalium.oru	LIAN .
O 2501/2007 Agreste, Nord-Pas-do-Calais	1277-5134 DRAF		1996 20	2005	44 (ms) 2005)	006	2		e	www.arrords.arrivally.co.	
	1621-3874 Journaux officiels	bireostrie		_	6 (novdéa, 2005)	000	٤		an kare	www.leditance.com/fr	ino.
	1282-7924 Direction des Journaux officides, Service de Mnformation et des roi	o trimostriel	1965	2006 11	4 (4e bim. 2005)	900	2	렒		st verog constitution www	ONE
28/01/2007 Bullotin official Ministers du Travail	The BACK TOUR		_	9	4000	Torris and the state of the sta	-		en Hone	финм рединатерия	
1902/2007 Reference	1287-6957 DRJS Rione-Aloes	annuel	1988	2004 6	2002	000	2 8	SC #	en fare	Wew, agreete agriculture, goung	non
28/02/2007 Prevention Parts, 1996)	1243-7123 ADHYS	mimostriol	-	₩	35 (2005)	006	5 5	DEP online	ou ligno	II ATOS SERVICIOS CAPA CAMBANIA I FRANCE	
	1703-4896 Eller	rimostriol	-1	9 9002	26 (soft 2005)	009	9	SCI	SCT on ligno		
2802/2007 Gense service	1708-0301 Organization generale des consommateurs	himeshiel	2004	2005	24 (2e tulm, 2005)	006	2 8	8	on ligne	MANA DE COLORES	
CONTROL OF THE PROPERTY OF THE	1,004-0-1,00 Estates Messource ingenetite de l'incernation	Billicomia	+	cm o	(coops same) /4	0%	2 1	2	cn ligne		
CONTRACT COLONIA SECTION DE LA COLONIA DE COLONIA DECENDA DE COLONIA DE COLON	COOPERS OF STREET	Solk is	-	9	crinzaionz uncurs	000	2		en ligno	Www.vaidoksalf	
anauto cartena	1/254-2562 Editions Chimideles		_	2002		009	2		an ligne	www.lulu.com	-
	1763-9093 Syndical national des hökelliets, rentauntieurs, carletiens et traiteur	Ě	_	2006	a 1,276 (sept, 2005)		õ		counfel		
Q2/33/2007 Global prace freedom world lour	1914-0961 Reporters without borders	anune	+	7007	2004	000	ę		ou ligno	# promon	
	1629-1468 DAC	mensuel	-	4	31 (dec.2005)	950	8	à	2		
	1631-2635 CDOS 53	troks foks parr ar	+	-		760	8	¥.	7		
DSICOROGO Tinfo Capitale	1270-4536 Chambre de commerce și d'Industrie de Paris	samestriel	+	2005 12		300	6		on land	www.cclo75.fr > consulter > le torifleire en chiffres	
09/03/2007 Eludes de la DRE (les)	1633-1559 DRE Pleardle	mensuel		_	4D (9 janv. 2007)	. 000	2		countiel		
12/03/3007 Washington flo	1276-3516 Ambassade américaina	quotidlen	-	3006	28 (12 few. 2007)	300	8	da DEb	en ligne	http://palnio.gov/unsinfo/products/washfile.html	모
18/03/2006 Lettre dos événements du toutisme et de la gastronomie	1766-9529 Cilé	anne	+	2006 2	mars-06	3000	မှု	꿅	an land		
19/09/2007 Tablesu do boxd	1788-028X Observatoire départemental de l'eingloi et de la formation du Ter	ritoke de Bellort	2000	2002 88888	15 (sept. 2005)	006	2	Ð	on igno		
390357307 Gabeco	1295-9030 Piems Le Roy	trimestrial	1998	2003 5	33 (mars 2004)	3000	8	8	er ligne		
17704/2007 Lettre AchalGrenoblo	1952-6970 CCI Grenoble	semestries	2002	1 2006	dec. 2006	300	9	GE	countiel		messageris
17/04/2007 Steves	1279-5964 Sciences po Grenoble	3×an	1996	2007	Mn-04	0000	ā	뫎	counted		messagerie
18/04/2007 Lettre aux clubs	1526-9083 Foderation française de gelf	3×m	1997	2007	57 (fevt. 2007)	700	5	∄	adrosso de toliteturgement envente par mail		messaggio
200542007 Hyghov lette d'information	Centra regional drincovation et de translet de technologie en 1165-7693. Inveriene alimentalire	3×82	1991	2006	50 (févr. 2006)	009	2	SCT	pourriel		messagerie
25004/2007 Bitalegraphie prospective	0291,123X Association Futuribles international	trensuel	1982	2006 24	44 (déc. 2006)	0	2	ă			
28/04/2007 Cahlers do l'OCBF	1770-4014 Office de cográfination bancaire et financière	mensuel	2004	2008	dec. 2006	Page 1 300	ē	430	countiel		
CO 27/04/2007 Vie scelatre, perents, 91	0151-2161 Conzeil départemental des parents d'élèves	trimestriel	1976 20	9008	S7 (sept. 2006)	300	5	뫒	on ligne		
040000001 Ligine d'ombre	1778-8382   Sockite des Amis de Gérald Herve	Monacom	2006 20	2006	1 flanv. 2006)	RDD	E	-	an live	_	

# Annexe 4 : Schéma présentant le fonctionnement de SPAR





# Annexe 5 : Extrait du panier de sélection du DDL validé par le DLN

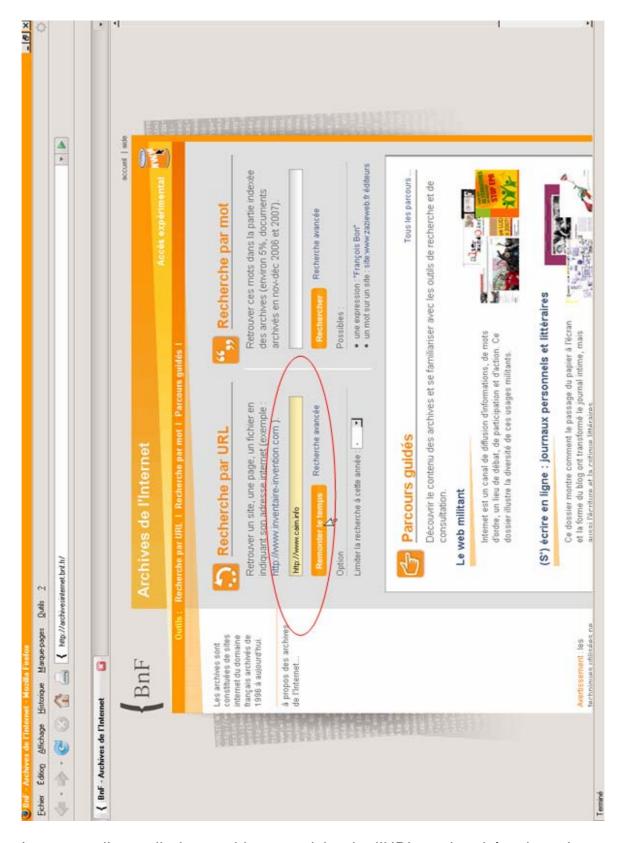
Panier	Panier URL de départ	Profondeur	Fréquence	Budget	Budget Large/ciblée	URL à inclure	Typologie
DDL	http://agriculture.godv.fr	domaine	1 fois par an	gros	ciblée		0
TOO	http://www.aisne.pref.gouv.fr	hôte	1 fois par an	moyen	ciblée	http://www.aisne.pref.gouv.fr/2007/i_p.php3?n p=79	Périodiques migrés
DDL	http://www.apce.com	hôte	1 fois par an	gros	ciblée		0
DDL	http://www.archives.premier-ministre.gouv.fr	hôte	1 fois par an	gros	ciblée		PO
DDL	http://www.assemblee-nationale.fr	hôte	1 fois par an	gros	ciblée		PO
DDL	http://recherche2.assemblee-nationale.fr	hôte	1 fois par an	gros	ciblée		PO
DDL	http://www.banque-france.fr	domaine	1 fois par an	gros	ciblée		Périodiques migrès
DDL	http://www.boc.sga.defense.gouv.fr	hôte	1 fois par an	moyen	ciblée		Périodiques migrés
DDL	http://www.budget.gouv.fr	domaine	1 fois par an	gros	ciblée		PO
DDL	http://www.cg16.fr	domaine	1 fois par an	gros	ciblée		PO
DDL	http://www.cg29.fr	domaine	1 fois par an	gros	ciblée		PO
DDL	http://www.cg74.fr	domaine	1 fois par an	gros	ciblée		PO
DDL	http://www.cndp.fr	hôte	1 fois par an	gros	ciblée		PO
DDL	http://www.cr-bourgogne.fr/	domaine	1 fois par an	gros	ciblée	http://www.cr- bourgogne.fr/.773.objet.publication.publication thems.2007-2013 http://www.cr- bourgogne.fr/?&objet=publication&categorie=7 73&publication thema=2001-2007 http://www.cr-bourgogne.fr/Avis-Mandature- 1985-2001.773,2352 http://www.cr- bourgogne.fr/Avis-Mandature- http://www.cr- bourgogne.fr/Publications.773,2350	PO .

### Annexe 6 : Le contenu du « Panier actualités »

Thème	Nom	Url	priorité
PQN	Aujourd'hui	http://www.aujourdhui-en-france.fr	1
PQN	Croix (La)	http://www.la-croix.com/	1
PQN	Echos (Les)	http://www.lesechos.fr/	1
PQN	Equipe (L')	http://www.lequipe.fr/	1
PQN	Figaro	http://www.lefigaro.fr/	1
PQN	France Soir	http://www.francesoir.fr/	1
PQN	Humanité	http://www.humanite.fr/	1
PQN	Libération	http://www.liberation.fr/	1
PQN	Monde (Le)	http://www.lemonde.fr/	1
PQN	Présent	http://www.present.fr/	1
PQN	Tribune (La)	http://www.latribune.fr	1
PQR	Bien Public (Le)	http://www.bienpublic.com/	3
PQR	Charente Libre (La)	http://www.charentelibre.com	3
PQR	Corse Matin	http://www.corsematin.com/	3
PQR	Courrier Picard (Le)	http://www.courrier-picard.fr/	3
PQR	Dauphiné Libéré (Le)	http://www.ledauphine.com/	<del>                                     </del>
PQR	Dépêche de Tahiti (La)	http://www.ledauphine.com/	1 3
PQR		http://www.ladepeche.fr/	2 3 3 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3
PQR	Dépêche du Midi (La)  Dernières Nouvelles d'Alsace (Les)	http://www.dadepeche.n/	2
		http://www.estrepublicain.fr/	2
PQR	Est Républicain (L')		3
PQR	France Antilles Guadeloupe	http://www.guadeloupe.franceantille	3
PQR	France Antilles Martinique	http://www.martinique.franceantilles	3
PQR	France Guyane	http://www.franceguyane.fr/	3
PQR	Journal de l'île de la Réunion (Le)	http://www.clicanoo.re/	3
PQR	Journal du Centre (Le)	http://www.lejdc.fr	3
PQR	Marseillaise (La)	http://www.journal-lamarseillaise.co	3
PQR	Midi Libre (Le)	http://www.midilibre.com/	3
PQR	Montagne (La)	http://www.lamontagne.fr/	3
PQR	Nice Matin	http://www.nicematin.com/	3
PQR	Nouvelle République (La)	http://www.lanouvellerepublique.fr/	3
PQR	Nouvelles Calédoniennes (Les)	http://www.info.lnc.nc/	3
PQR	Ouest France	http://www.ouest-france.fr/	
PQR	Paris Normandie	http://www.paris-normandie.fr	3
PQR	Populaire du Centre (Le)	http://www.lepopulaire.fr/	
PQR	Progrès de Lyon (Le)	http://www.leprogres.fr	2
PQR	Provence (La)	http://www.laprovence.com/	2
PQR	Quotidien de la Réunion (Le)	http://www.lequotidien.re/	3
PQR	Républicain Lorrain (Le)	http://www.republicain-lorrain.fr	3
PQR	Sud-Ouest	http://www.sudouest.fr/	2
PQR	Télégramme de Brest (Le)	http://www.letelegramme.com/	3 3 2 2 2 3
PQR	Union (L')	http://www.lunion.presse.fr/	3
PQR	Voix du Nord (La)	http://www.lavoixdunord.fr/	1
PQR	Yonne Républicaine (L')	http://www.lyonne.fr	3
Pmag	Express (L')	http://www.lexpress.fr/	1
Pmag	Marianne	http://www.marianne2.fr/	1
	Nouvel Observateur	http://www.nouvelobs.com/	1
Pmag Pmag		http://www.lepoint.fr/	1
Pmag Pmag	Point (Le)		· 1
Pmag	Journal du dimanche (Le)	http://www.lejdd.fr/	2
Pmag	Expansion (L')	http://www.lexpansion.com/	
Agence	Reuters	http://fr.reuters.com/	1
PQN	20 minutes	http://www.20minutes.fr/	1
PQN	Metro	http://www.metrofrance.com	1
Buzz	Le Post	http://www.lepost.fr/	2 2 2
Portail	Yahoo	http://fr.news.yahoo.com/	2
Portail	24 24 actu	http://www.2424actu.fr/	2

Portail	MSN	http://news.fr.msn.com/m6-actualite	2
Portail	Orange	http://actu.orange.fr/	2
Buzz	Buzz actu	http://www.buzz-actu.com/	. 3
Portail	Info Politique	http://infopolitique.free.fr/	3
Buzz	Les Indiscrets	http://www.lesindiscrets.com/	2
Portail	Internaute (L')	http://www.linternaute.com/actualite	2
Portail	Laposte	http://www.laposte.net/thematique/a	2
Portail	Voilà	http://www.voila.fr/	2
pure player	Rue 89	http://www.rue89.com/	1
pure player	AgoraVox	http://www.agoravox.fr/	1
Portail	Google	http://news.google.fr/	2
pure player	Slate	http://www.slate.fr	1
pure player	Médiapart (le club)	http://www.mediapart.fr/club	1
pure player	Bakchich	http://www.bakchich.info/	1
Pmag	Télérama	http://www.telerama.fr	2
Pmag	Les Inrockuptibles	http://www.lesinrocks.com	2
infonet	Numerama	http://www.numerama.com	2
infonet	Ecrans	http://www.ecrans.fr	2
infonet	Wikio	http://www.wikio.fr	2
infonet	PCINpact	http://www.pcinpact.com	2

## Annexe 7 : Exemple d'une consultation dans les Archives de l'internet (Cairn)



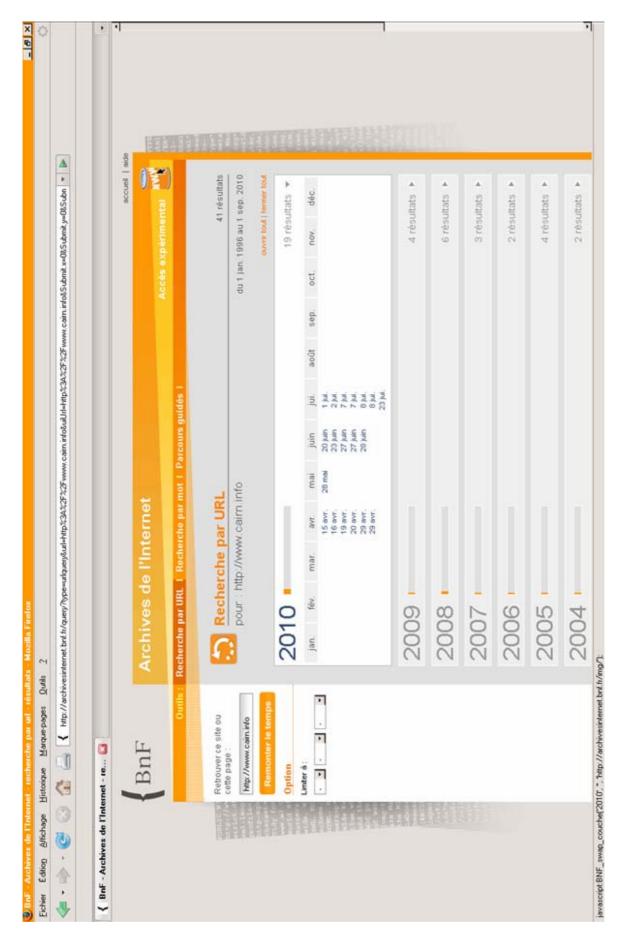
La page d'accueil des archives : saisie de l'URL recherchée dans la zone « Recherche par URL » « remontons le temps »...





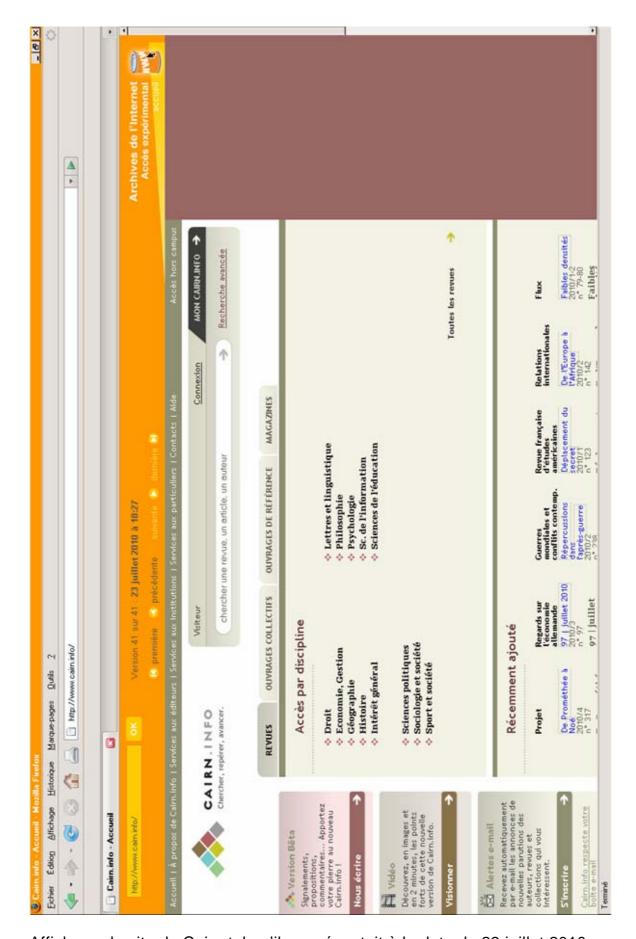
Accès au calendrier permettant de choisir le millésime de la collecte





Accès au calendrier permettant de choisir la date précise de la collecte



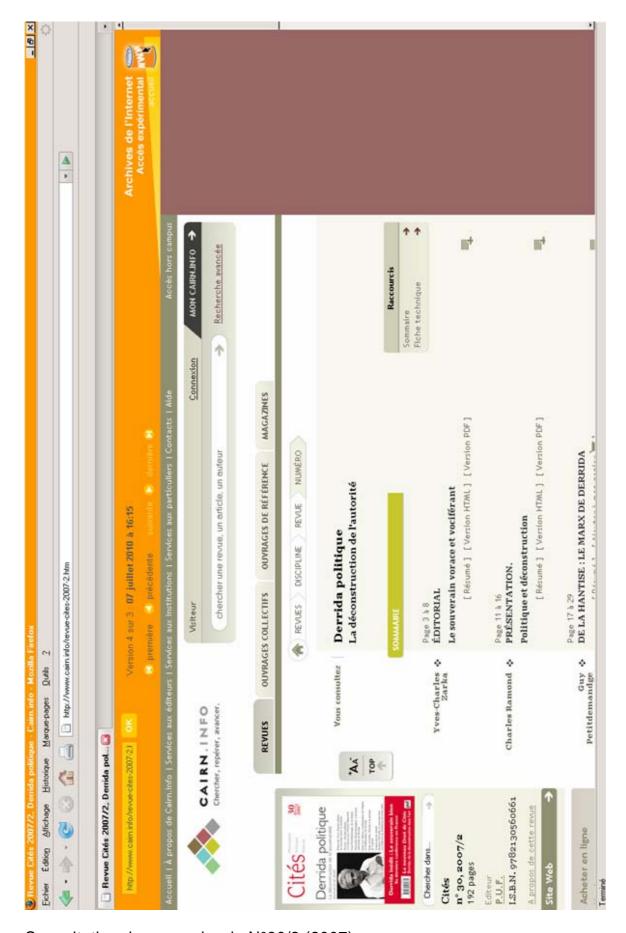


Affichage du site de Cairn tel qu'il se présentait à la date du 23 juillet 2010



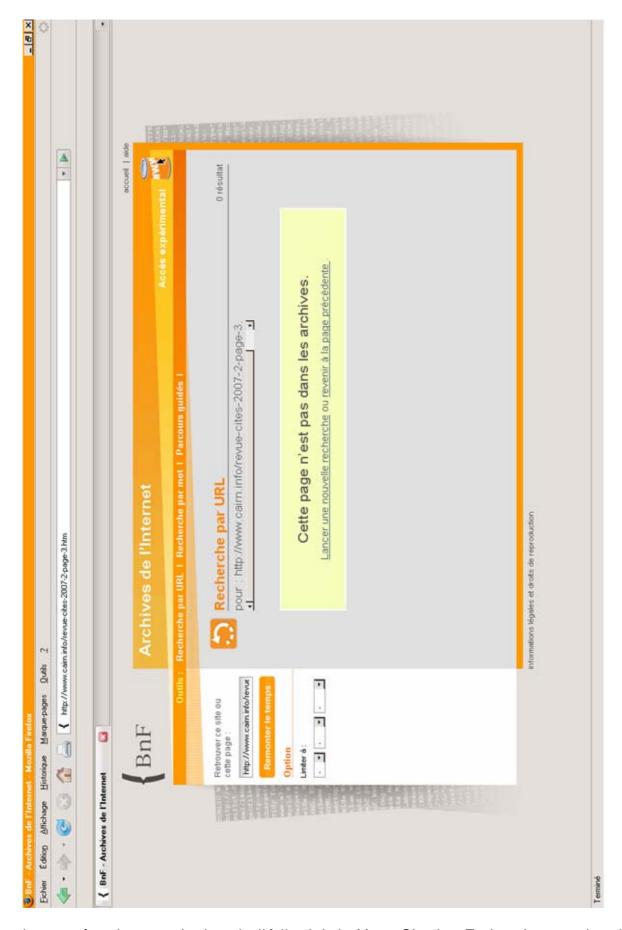


Consultation de la revue Cités au sein de cette archive



Consultation du sommaire du N°30/2 (2007)





La requête de consultation de l'éditorial de Yves-Charles Zarka n'a pas abouti. La collecte n'est pas descendue au niveau de l'article dans cette archive



### Annexe 8 : Courrier type relatif aux publications migrées



direction des Services et des réseaux Quai François-Mauriac 75706 Paris Cedex 13 Téléphone 33(0) 1 53 79 88 83 - télécopie 33 (0) 1 53 79 85 56 nº Sirut 18004625200177 - code APE 9101Z - nº didentification TVA FR 88 180 046 252

Paris, le

Société Adresse 1 Adresse 2 Code postal VILLE

A l'attention de

GDP 022 Publication migrée en ligne ou diffusée par courriel

Monsieur,

département du Dépôt légal service de Gestion des périodiques

tél: 33 (0) 1 53 79 54 38 marie.guibert@bnf.fr

réf: BnF-ADM-2008-008281-01

affaire sulvie par : Marie GUIBERT

Nous avons bien reçu votre courrier daté du ..... par lequel vous nous informez que la publication intitulée ..... cesse de paraître sous la forme imprimée.

Depuis la parution de la loi du 1er soût 2006 relative aux droits d'auteur et aux droits voisins dans la société de l'information (DADVSI), le dépôt légal de l'Internet est régi par le Titre III du Code du patrimoine. Ce texte habilite la BnF à collecter régulièrement les sites Web du domaine français et à conserver ces données. Cette collecte se fait de manière automatisée par le passage d'un robot.

Contrairement au dépôt légal traditionnel des publications imprimées, le dépôt légal de l'Internet n'implique pas de demarche active de votre part à moins que tout ou partie de votre site soit inaccessible au robot de capture pour des raisons techniques (base de données, contenu protégé par mot de passe, formulaire d'accès...) ou économiques (contenu payant, abonnement ...). Dans ce cas, la BnF peut être amenée à prendre contact avec vous au cas par cas.

Pour plus de renseignements, vous pouvez consulter les pages de notre site consacrées au dépôt légal du web : www.bnf.fr > Pour les professionnels > Dépôt légal > Sites Web.

L'attribution d'un ISSN à un périodique en ligne ne s'effectue pas automatiquement dans le cadre du dépôt légal de l'Internet. Il n'est pas obligatoire mais, si vous le souhaitez, vous pouvez en faire la demande en téléchargeant le formulaire à partir du site Web de la BnF :

www.bnf.fr > Pour les professionnels > ISSN, ISBN, autres numéros > S'informer et obtenir un ISSN > Obtenir un ISSN > Document en ligne ou diffusé par courrier électronique > Formulaire de demande d'ISSN pour les ressources en ligne.

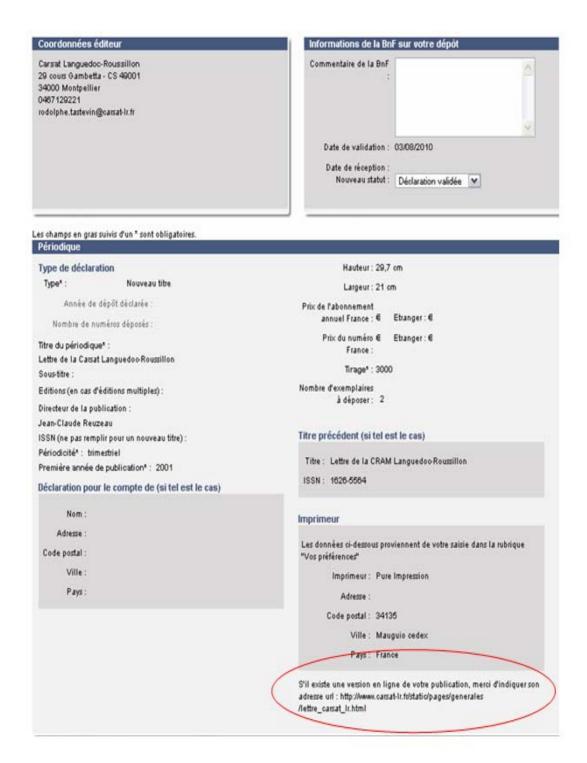
Si toutefois votre publication était à nouveau diffusée sur support papier, vous seriez tenu d'en faire le dépôt, même en cas de coexistence avec une version électronique.

Je reste à votre disposition pour tout renseignement complémentaire. Je vous prie d'agréer, Monsieur, l'expression de ma considération distinguée.

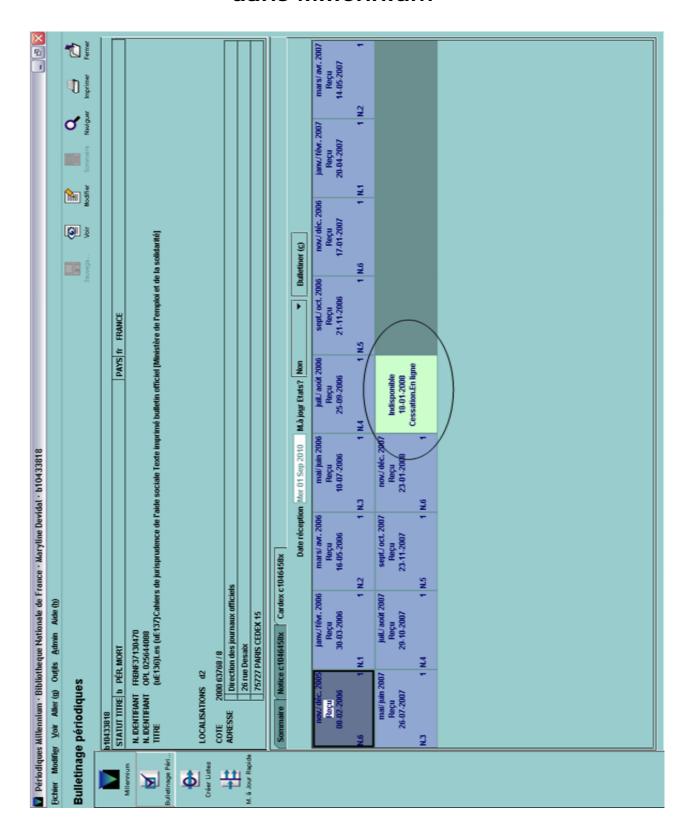
> Alix BRUYS chef de service



#### Annexe 9 : Formulaire de déclaration extranet



# Annexe 10 : le cardex d'un périodique migré dans Millennium

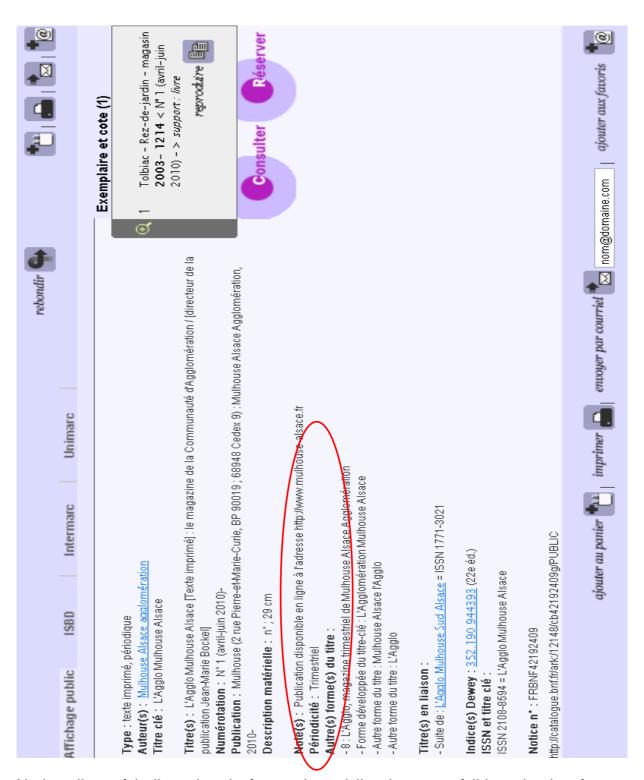


# Annexe 11 : Formulaire de signalement d'un périodique migré

		Bibliothèque na	tionale de France
		•	vices et des réseaux
		Département du	
		Service gestion d	
		qual François Mau téléphone	riac 75706 Paris cedex 13 33 [0] 1 53 79 42 59
		rélécople	33   0) 1 53 79 85 50
		mél	Maryline.Cevidal@bn1.fr
	affaire suivie par : Maryline Devidal		
Objet	Signalement d'un périodique vers une édition en ligne. <sup>1</sup>	e imprimé reçuautitre d	u dépôt légal migrant
<u>Date</u> :			
DBIC .			
	<u>in du périodique :</u>		
ISSN (éd.			
Imprimée)			
Auteur(s) ou collectiv	ritá		
éditrice	nue		
Titre :sous	-titre		
Editeur, an	née		
20.030.   2			
URL à Jaqu	elle		
peut être			
consulté le			
périodique			
URL du site	•		
hébergeur			
<u>Origine de la </u>	a de mande :		
Départeme			
service			
Nom, prén	om		
Téléphone	ou		
courriel			
Inches Inches		- 4-1	
- je souhaite ———	e connaître le résultat de cette	e réclamation : oui	non
		à l'étranger et diffisés en Fran	

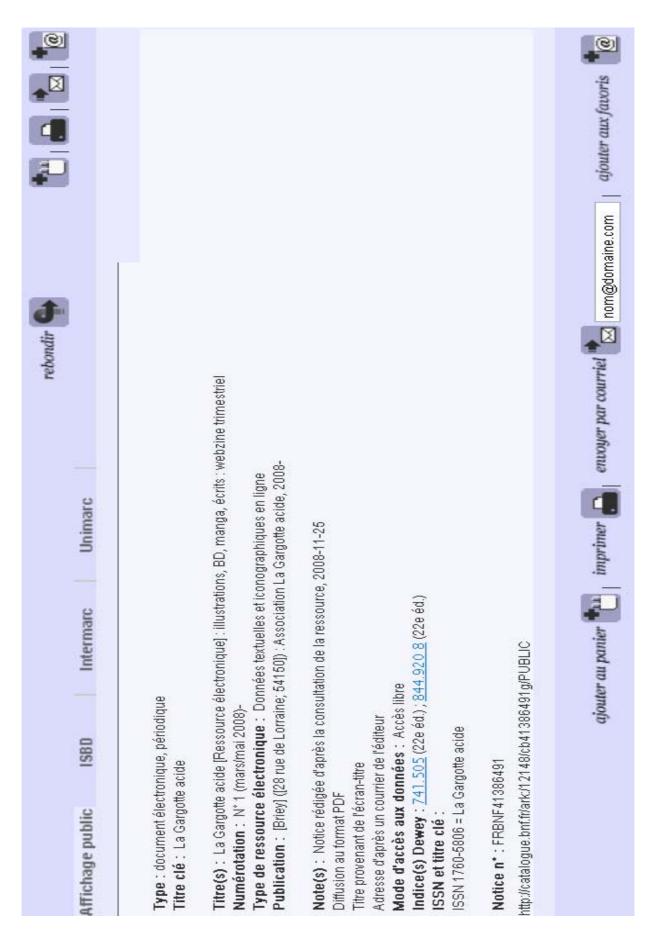


### Annexe 12 : Exemples de notices mentionnant une édition en ligne dans le Catalogue général de la BnF



Notice d'un périodique imprimé pour lequel il existe une édition simultanée en ligne. L'URL a été saisie en zone de note publique. L'édition en ligne n'a pas été cataloguée





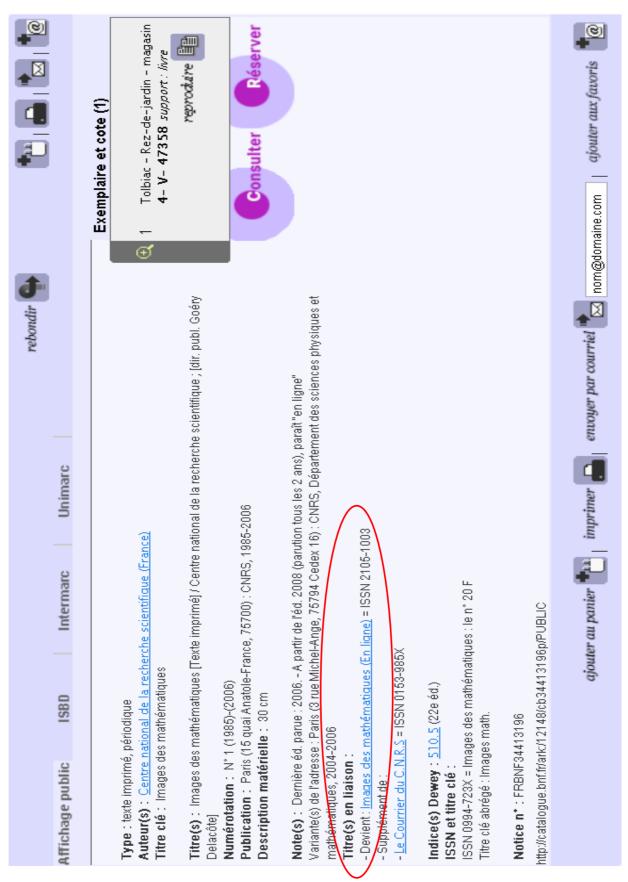
Notice d'un périodique né numérique. L'URL n'est pas visible en affichage public bien qu'elle ait été saisie (cf vue suivante)



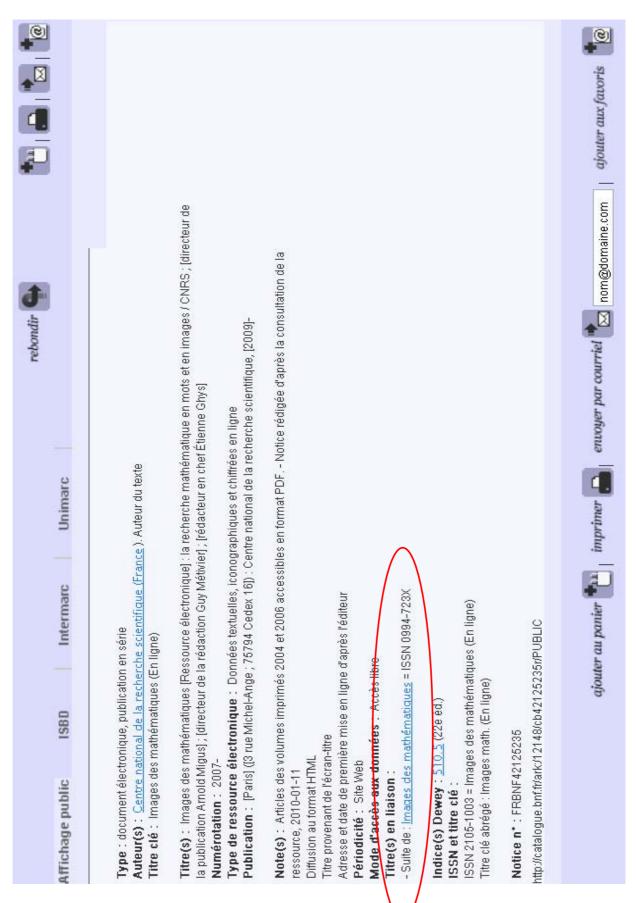


La même notice au format intermarc. L'URL a été saisie en zone 856, zone non visible en l'affichage public





Notice de l'édition imprimée d'un périodique ayant migré en ligne : les zones de numérotation et de la date de publication sont closes. Un lien a été fait vers la notice de l'édition en ligne



Notice de l'édition en ligne faisant suite. Ici l'URL N'est pas visible car elle figure en zone 856 du format intermarc

## Annexe 13 : Une fiche de prospection ayant reçu une réponse Non : en ligne

