

10 Avril 2007

A L'ATTENTION DE :
BIBLIOTHÈQUE NATIONALE DE FRANCE
Quai François Mauriac
75706 Paris cedex 13

ETUDE EN VUE DE L'ELABORATION D'UN
MODELE ECONOMIQUE DE PARTICIPATION DES EDITEURS
A LA BIBLIOTHEQUE NUMERIQUE EUROPEENNE (*EUROPEANA*)

1	<i>Problématique de l'étude et méthodologie</i>	2
1.1	Problématique	2
1.2	Méthodologie	3
2	<i>Analyse de l'environnement : la chaîne du livre numérique et les modèles économiques existants</i>	5
2.1	Les livres numériques : concepts, usages, valeur ajoutée	5
2.2	La « chaîne du livre numérique »	8
2.3	Typologie des modèles économiques	12
2.3.1	La commercialisation des contenus numériques sous droits	12
2.3.2	Les modèles économiques de librairie numérique	15
2.3.3	Les modèles économiques de bibliothèque numérique	17
2.3.4	Vers des modèles hybrides ?	19
2.4	Les mesures techniques de protection des droits d'auteur ou « DRM »	21
2.4.1	L'utilité des DRM ?	21
2.4.2	Les DRM pour la lecture en ligne	22
2.4.3	Les DRM pour le téléchargement	24
2.5	Repérages : quelques acteurs de la distribution de livres numériques	25
2.5.1	Les librairies de livres numériques	26
2.5.2	Les offres de bibliothèques numériques par les agrégateurs	28
2.5.3	Les offres numériques directes des éditeurs	30
3	<i>Recommandation : le modèle économique proposé</i>	31
3.1	Les fonctionnalités du site Europeana	32
3.2	Les principaux objectifs et réquisits de la BnF pour Europeana	38
3.3	Les principaux objectifs et réquisits des éditeurs pour participer à Europeana	39

3.4	Analyse des modèles économiques envisageables et recommandation	40
3.5	Le modèle économique recommandé : quels partenaires ?	44
3.5.1	Les opérateurs de diffusion agréés	44
3.5.2	La place des librairies dans le modèle économique d'Europeana	47
3.5.3	Le cas particulier des lecteurs handicapés	49
3.5.4	L'accord des auteurs et des ayants droit	51
3.6	Le parcours type des usagers	53
3.7	Les contrats de partenariat avec les diffuseurs agréés et les transferts de données	57
3.7.1	Normalisation et transfert des métadonnées	58
3.7.2	Normalisation et transfert des fichiers texte	61
3.7.3	Protocole d'authentification partagé	62
4	La numérisation des contenus et la politique documentaire	64
4.1	Processus et coûts de numérisation	64
4.1.1	Numérisation stricto sensu	64
4.1.2	Conversion ou adaptation numérique	65
4.1.3	Avantages comparatifs	66
4.2	Choix de numérisation et politique documentaire	67
4.3	Simulation économique	70
5	Entretiens, acceptabilité et variantes	75
5.1	Entretiens réalisés	75
5.2	Acceptabilité	77
5.3	Variantes	78
6	Mise en œuvre et compléments d'étude recommandés	79

1 Problématique de l'étude et méthodologie

1.1 Problématique

La Bibliothèque nationale de France (« BnF ») s'est vu confier la responsabilité opérationnelle de la contribution française à *Europeana*, projet de bibliothèque numérique européenne, qui vise à mettre en ligne, à un horizon de cinq ans, plusieurs millions d'ouvrages représentatifs de l'identité, de l'histoire et de la culture européenne. Elle s'appuie pour ce faire sur l'expérience acquise grâce à sa propre bibliothèque numérique, *Gallica*, qui propose aux internautes plus de 80 000 ouvrages, libres de droits, gratuitement accessibles en ligne et téléchargeables. La bibliothèque numérique européenne a également vocation à proposer aux internautes des documents récents encore soumis au droit d'auteur, à côté de documents libres de droits et patrimoniaux.

La présentation sur Internet de documents sous droits ne peut être envisagée que dans le cadre d'un dispositif assurant une juste rémunération des ayants droit et d'une manière

générale un respect du code de la propriété intellectuelle. Les ayants droit étant, dans l'économie du livre, représentés pour l'essentiel par des éditeurs avec lesquels des contrats d'auteur ont été établis, la BnF a confié à la société Numilog.com (« Numilog ») la présente étude en vue d'élaborer un modèle économique de participation des éditeurs à *Europeana*.

Le projet de bibliothèque numérique européenne représente un des plus ambitieux projets de mise à disposition du public de documents numériques jamais imaginés, dans un contexte international où plusieurs projets privés de numérisation ont été lancés, pour certains sans les mêmes objectifs de service public, d'organisation intelligente des corpus de documents et de respect des droits d'auteur et des acteurs de la chaîne du livre.

La mise au point d'un modèle économique cohérent et pertinent permettant à *Europeana* d'inclure des documents sous droits représente un facteur clé de succès de ce projet. Pour autant, la complexité de la question ne doit pas être sous-estimée, car d'une part elle met en jeu des groupes d'acteurs aux intérêts ex ante différents, et d'autre part elle fait intervenir un ensemble de contraintes économiques, fonctionnelles, techniques et juridiques sans la prise en compte desquelles tout modèle économique théorique serait vain.

A vocation européenne, le démarrage d'*Europeana* est piloté par la BnF qui s'adressera en priorité à des éditeurs français pour y inclure de premiers contenus sous droits selon un modèle destiné à servir d'exemple de pratique réussie pour des bibliothèques nationales et des éditeurs d'autres pays européens. L'étude a donc été menée en priorité à travers des discussions avec les acteurs français de la chaîne du livre, mais une première investigation de l'acceptabilité du modèle proposé par des acteurs d'autres pays européens a été conduite.

1.2 Méthodologie

L'analyse menée vise à envisager plusieurs schémas pour construire un modèle économique pour la présentation de documents sous droits dans le cadre d'*Europeana*.

La méthode générale adoptée consiste à identifier le ou les modèles économiques qui permettraient de concilier deux gammes d'objectifs correspondant aux deux groupes en discussion pour initier le projet :

- les objectifs et contraintes de la BnF pour *Europeana* tels qu'ils sont définis dans le cahier des charges de l'étude et tels qu'ils ont été précisés par les responsables du projet au cours de l'étude ;
- les objectifs et contraintes des éditeurs pour participer à *Europeana* en y « apportant » (en un sens à définir) des ouvrages dont ils détiennent les droits d'exploitation tels qu'ils ressortent d'entretiens avec plusieurs éditeurs et représentants d'éditeurs.

A ces deux gammes d'objectifs peuvent être subordonnés les objectifs connexes suivants :
- la faisabilité technique des schémas envisagés et en particulier les garanties offertes par les mesures de protection et de calcul de la rémunération des ayants droit ;

- une totale compatibilité avec l'environnement légal et réglementaire, notamment dans le cadre de la loi DADVSI en France et plus généralement de la directive européenne que cette loi a transposé et des dispositions régissant la propriété intellectuelle ;
- l'intégration, « tant que faire se peut », dans les schémas optimaux de l'ensemble des groupes d'acteurs de la chaîne du livre, c'est-à-dire la prise en compte de leurs rôles et intérêts respectifs ;
- la prise en compte des expériences et des exemples en France et à l'étranger, notamment le succès rencontré par différents modèles économiques de diffusion de contenus numériques ;
- l'organisation générale actuelle et potentielle de la filière professionnelle de l'édition numérique ;
- les coûts éventuels d'investissement et de fonctionnement à engager.

L'étude s'appuie d'abord sur l'expérience de Numilog sur le marché des livres numériques et dans la mise au point de solutions techniques, juridiques et commerciales adaptées à ce marché et validées par ses acteurs, notamment des modèles de tarification et des solutions de DRM associées. Des typologies permettant de caractériser les différents processus à l'œuvre dans la filière du livre numérique et de dresser un panorama des catégories d'acteurs, de modèles économiques et techniques à l'œuvre dans ce secteur sont proposées. Ces typologies ont pour but de servir d'outils conceptuels et cartographiques pour produire des réponses pertinentes à la question centrale de l'étude, en éliminant les réponses trop intuitives ou contradictoires. Une attention particulière sera portée au rôle spécifique que peut ou devrait jouer une institution publique telle qu'*Europeana* dans ce panorama, en prenant soin de différencier son rôle de celui d'acteurs commerciaux.

En ce qui concerne l'analyse de l'environnement et des modèles possibles, l'exhaustivité de l'analyse théorique est systématiquement privilégiée sur celle des comptes rendus pointillistes d'expériences concrètes, conçus comme de simples illustrations de cette analyse ou destinés à mettre en évidence les catégories les plus ancrées dans le réel. **Il ne s'agit pas de proposer à *Europeana* d'imiter telle ou telle expérience existant ailleurs, mais de proposer un modèle nouveau issu d'une réflexion complète sur les modèles logiquement possibles.**

L'étude s'appuie par ailleurs sur :

- la prise de connaissance des études déjà effectuées et des données déjà rassemblées par le groupe de travail BnF/SNE, notamment dans le domaine des fonctionnalités attendues – des processus de numérisation, des solutions de lecture et de recherche adoptées, des normes techniques retenues pour les fichiers et les métadonnées ; le « Livre Blanc du Comité de Pilotage pour une Bibliothèque Numérique Européenne » établi en janvier 2006, les comptes rendus du groupe de travail BnF/SNE et le prototype de site www.europeana.eu ouvert depuis le 23 mars 2007 sont en particulier des sources d'information capitales ;
- la consultation et l'analyse d'un ensemble de documents et de publications portant sur des expériences passées ou en cours ou sur l'économie de l'édition numérique en général ;
- une série d'entretiens avec des responsables de BnF, dans les domaines techniques, juridiques et documentaires ;
- une série d'entretiens avec différents acteurs de la chaîne du livre, en France et à l'étranger, éditeurs, bibliothécaires, libraires, experts de l'économie numérique. Ces entre-

tiens ont permis de tester l'acceptabilité des schémas envisagés, de les amender ou de les compléter par de nouvelles suggestions et d'orienter l'étude en vue d'une recommandation.

L'étude est livrée sous la forme :

- d'un premier compte rendu oral devant le groupe de travail BnF/SNE, qui a eu lieu le 20 février 2007, afin de présenter les premières hypothèses, de s'assurer de la bonne orientation de l'étude et d'intégrer des remarques ou suggestions du groupe de travail ;
- du présent document écrit reprenant les données rassemblées pour l'étude, l'analyse, les différents schémas envisagés et la recommandation de Numilog ;
- d'un deuxième compte rendu oral devant le groupe de travail BnF/SNE, qui aura eu lieu à la fin du mois d'avril 2007.

2 Analyse de l'environnement : la chaîne du livre numérique et les modèles économiques existants

2.1 Les livres numériques : concepts, usages, valeur ajoutée

Les **livres numériques**, ou « **ebooks** », sont des livres dématérialisés, qui peuvent être transportés et lus par le moyen d'un fichier informatique. Dans la plupart des cas aujourd'hui, le moyen de transport le plus commode se trouve être Internet.

Trois usages principaux des livres numériques sont alors possibles :

1. La **recherche plein texte** dans une base de livres, en utilisant des moteurs de recherche dans les contenus de tous les livres de cette base.
2. La **consultation en ligne** de pages des livres, soit dans des extraits, correspondant notamment aux résultats de la recherche plein texte, soit dans le texte intégral des ouvrages.
3. La **lecture hors ligne** du livre après téléchargement.

Lorsqu'ils sont lus hors ligne, la livraison des livres numériques **n'est pas effectuée nécessairement par Internet**. Internet n'est a priori qu'un outil parmi d'autres pour transporter les ebooks, qui pourraient aussi être transportés sur les réseaux téléphoniques hauts débits, ou simplement sur des CD-ROM ou sur des cartes mémoires amovibles, comme les « SD Cards », qui peuvent être vendus dans des magasins physiques. Internet, surtout, n'est qu'un outil parmi d'autres pour les lire. Il est donc faux d'une manière générale d'assimiler les livres numériques aux « livres sur Internet ».

Un point en particulier est important à souligner : **les livres numériques ne sont pas des pages Internet**, et n'ont **rien à voir avec les blogs**, qui reposent sur une mise à jour permanente et une interaction en continu d'un auteur de blog et d'internautes et qui ne font l'objet d'aucune sélection éditoriale. Les livres numériques sont d'authentiques livres, versions numériques d'ouvrages qui sont en général par ailleurs édités et diffusés sous forme imprimée. Bien sûr ils peuvent être découpés par chapitre, et comporter des « plus » par rapport à la version imprimée, comme des liens hypertexte ou des vidéos incrustées.

Certains ebooks peuvent n'exister qu'en version numérique. Dans tous les cas ils demeurent des livres, avec un début et une fin malgré quelques expériences très marginales encore de littératures hypertextuelles, en tous cas des œuvres cohérentes et achevées d'un auteur ou d'un collectif d'auteurs et fruits d'un travail éditorial complet sur le fond et sur la forme. Ce point est très important, d'une part pour délimiter le champ des œuvres sous droits qui ont vocation à être incluses dans *Europeana*, d'autre part pour répondre aux interrogations de spécialistes du livre ou d'auteurs qui verraient dans le livre numérique une menace pour la forme « livre ».

Les **appareils de lecture** utilisés pour lire des ebooks, ou « **e-readers** » s'il s'agit d'appareils dédiés à cela, sont potentiellement très divers : ordinateurs (notamment pour la lecture en ligne), agendas électroniques (« PDA »), smartphones. Les Japonais lisent même des ebooks sur des consoles de jeux. Tout ce qui porte écran peut, avec toutefois un confort variable, être utilisé pour lire un livre numérique. D'après une étude réalisée en 2006 par l'International Digital Publishing Forum, **83% des lecteurs se disent satisfaits ou très satisfaits du confort de lecture** offert par les appareils de lecture mobile existants du type PDA. Prochainement, les écrans à base d'**encre électronique**, avec des résolutions de l'ordre de 300 dpi, amélioreront encore ce confort, avec une impression de netteté voisine de celle des livres imprimés dans une ambiance de lumière naturelle. Les appareils utilisant une technologie d'affichage LCD, comme les PDA, conserveront toutefois un avantage : le rétro-éclairage, qui permet de lire dans des ambiances de faible éclairage. En outre, si les ebooks sont lus en ligne ou téléchargés sur un ordinateur, les pages peuvent être imprimées par les lecteurs préférant encore la lecture sur papier, en fonction bien entendu des droits définis par les éditeurs.

Conçus ainsi, les livres numériques présentent de très nombreuses **valeurs ajoutées fonctionnelles** pour les lecteurs, qui les différencient fortement des livres imprimés :

- l'accès immédiat à distance 24H/24 dans le monde entier ;
- la possibilité d'effectuer des recherches plein texte dans le livre ;
- la possibilité d'annoter le livre, sans le détériorer dans le cadre d'un usage collectif en particulier en bibliothèque ;
- la possibilité d'adapter la taille des caractères à sa vue et à son écran ;
- la possibilité d'utiliser des logiciels de synthèse vocale et celle de faire défiler automatiquement le texte, qui représentent des fonctions essentielles respectivement pour les handicapés visuels et moteurs ;
- la possibilité d'imprimer facilement des pages du document, en fonction des droits conférés par les éditeurs ;
- la possibilité de bénéficier d'enrichissements hypertextuels, voire multimédias (son, vidéos, animations flash) ;
- s'ils sont téléchargés, la mobilité pour des quantités importantes de contenus, ouvrant la possibilité de disposer de « bibliothèques nomades » personnelles.

Une étude réalisée au début 2006 par le consortium de bibliothèques Couperin dans 11 Universités françaises ayant testé pendant six mois le prêt de livres numériques indique que **73% des usagers interrogés étaient satisfaits par ce service**.

Pour autant, le bénéfice de toutes ces fonctionnalités n'est pas automatique et dépend largement des choix éditoriaux (pour l'enrichissement des livres), des **formats de numérisation** et des **logiciels de lecture** adoptés. Ces logiciels de lecture, ou « **Readers** » sont en général différents selon que la lecture est faite « en ligne » ou « hors ligne ».

Pour la **lecture en ligne** on peut par exemple présenter la matrice suivante des grandes options techniques, en fonction du type de contenu sous-jacent et du type de logiciel affichant ce contenu :

Type de logiciel → Type de contenu ↓	Navigation Html simple	Html amélioré (Java scripts, Ajax)	Application Java, Flash
Images (ex : PDF image avec texte sous image)	1	2	3
Flux de texte (ex : PDF texte ou HTML)	4	5	6

Les logiciels de lecture peuvent être très variés au sein de chacune de ces catégories, chaque acteur de distribution de contenus ayant tendance à créer le sien avec les fonctionnalités et l'interface qu'il juge appropriées, sans qu'aucune standardisation ne soit apparue. Les fonctionnalités offertes sont en général plus complètes si on se déplace vers la droite dans le tableau des solutions de lecture en ligne : les solutions à base d'applications Java ou Flash sont plus complètes que celles de simples logiciels de lecture HTML, et se rapprochent de celles offertes par les logiciels de lecture hors ligne. La protection des fichiers présente en général un caractère plus robuste si le contenu diffusé est composé d'images que de texte.

Pour la **lecture hors ligne**, la différenciation se fait entre un nombre moins grand de solutions de lecture, plus standardisées, dont les avantages comparatifs tiennent certes aux fonctionnalités, mais d'abord et surtout à leur compatibilité avec différents types d'appareils de lecture. On peut identifier aujourd'hui 4 formats de lecture les plus répandus, associés à 4 types de Reader, avec les compatibilités matérielles suivantes :

Type d'appareil → Format et Reader ↓	PC	Mac	Palm	Pocket PC	Smartphones (certains)
PDF Adobe Reader	X	X			
PRC Mobipocket Reader	X		X	X	X
PDB Palm Reader	X		X	X	X
LIT Microsoft Reader	X			X	X

A ces logiciels s'ajouteront sans doute prochainement des logiciels de lecture spécifiques pour des appareils de lecture dédiés utilisant la technologie d'encre électronique, comme le Sony Reader™ ; certains de ces appareils, comme le Illiad™ de Irex, étant toutefois vraisemblablement destinés à être compatibles avec certains des logiciels standards cités. Tous ces logiciels de lecture hors ligne présentent en général plus de fonctionnalités d'adaptation des textes à la vue du lecteur et à la taille de son écran que les logiciels de lecture en ligne.

Les usages des livres numériques

1. Recherche dans une base de textes
2. Lecture en ligne
3. Lecture hors ligne après téléchargement



2.2 La « chaîne du livre numérique »

Compte tenu de l'indépendance qu'ils permettent d'avoir, en termes de fabrication et de distribution, par rapport aux contraintes physiques liées à l'imprimé, il est légitime de se demander si l'édition numérique ne risque pas de bouleverser radicalement la chaîne traditionnelle du livre.

L'édition numérique semble ouvrir certaines marges de manœuvre pour que se déploient des stratégies de désintermédiation :

- les auteurs qui le souhaitent ou ne trouvent pas d'éditeur peuvent désormais plus facilement s'« autoéditer », en utilisant également les technologies d'impression numérique ;

- les éditeurs peuvent désormais plus facilement s'autodistribuer, à partir de leur propre site Internet (ils peuvent d'ailleurs le faire pour les livres imprimés, indépendamment des ebooks) ;
- les lecteurs eux-mêmes peuvent désormais plus facilement être sollicités au cours du processus de création d'un livre, voire intervenir dans son écriture. Ils pourraient aussi, en théorie, recréer un livre nouveau à partir des pages d'autres livres ou en créant un réseau de livres reliés hypertextuellement, selon la formule du « livre liquide » mise en avant par Francis Pisani.
(http://pisani.blog.lemonde.fr/pisani/2006/05/le_futur_du_liv.html).

Mais toutes ces possibilités ne devraient pas modifier profondément le rôle dominant des acteurs traditionnels de la chaîne du livre :

- **L'auteur** demeure bien sûr l'acteur premier et central de la création d'un livre, l'intervention des lecteurs dans le processus d'écriture demeurant expérimentale et plus que marginale pour la création littéraire, et son rôle en fin de parcours pour certains livres techniques (en informatique surtout pour le moment) n'enlevant rien au rôle dominant de l'auteur spécialiste. L'éventuel réassemblage d'un livre à partir de parties de livres ne constitue par une création à proprement parler et n'a probablement pas l'intérêt qu'on lui prête parfois, sauf dans des contextes particuliers existant déjà avec l'imprimé comme la création de dossiers de travaux dirigés.
- **La fonction d'édition** demeure essentielle à la sélection, à la mise en forme et à la diffusion de livres de qualité adaptés à un public et à un marché ; il n'y a pas besoin d'insister sur ce point tant il est évident que l'édition numérique de livres se perdrait, comme celle de musique, à se dissoudre dans une masse de textes non triés écrits par des millions d'« auteurs »... ce scénario n'est ni souhaitable ni crédible : encore une fois, les livres numériques ne sont pas des blogs.
- **Les fonctions de libraire et de diffuseur** garderont une place importante dans le circuit de l'édition numérique, pour les mêmes raisons, que dans le circuit de l'édition traditionnelle : elles simplifient la relation entre des lecteurs potentiels ou des médiateurs tels que les bibliothèques et un catalogue de livres, en offrant sur un même lieu et selon des modalités techniques et commerciales homogènes une offre multi-éditeurs par nature plus riche que celle de chaque éditeur séparément. En termes mathématiques, elle substitue un réseau de N relations (pour N lecteurs ou bibliothèques) à un graphe de NM relations (pour M éditeurs).
- **La fonction de bibliothèque** devrait rester longtemps importante, les livres numériques prenant leur place pour donner un sens au concept de « bibliothèque hybride ». Les bibliothèques numériques, émanant des bibliothèques physiques, ont en effet exactement le même rôle de médiation culturelle économique à jouer que pour les livres imprimés : celui de permettre à des lecteurs inscrits, qui représentent une communauté ayant des points communs (localisation, études...) d'accéder gratuitement, mais pour une durée limitée et sans appropriation, à des œuvres par ailleurs sélectionnées et organisées.

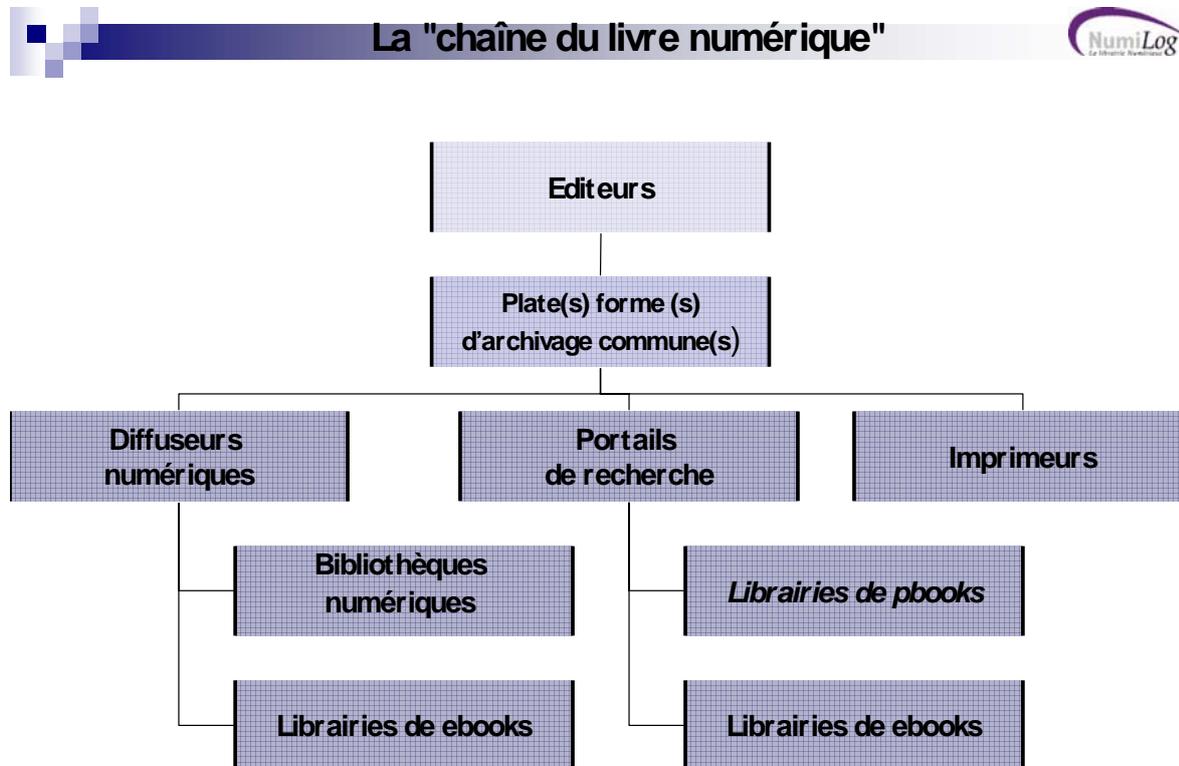
Si le rôle des acteurs traditionnels de la chaîne du livre est maintenu dans l'ensemble, il n'en reste pas moins que le numérique introduit d'une part une adaptation des processus métiers de fabrication et de distribution, et d'autre part l'apparition de certains acteurs et circuits économiques nouveaux, constituant ce qu'on peut appeler par analogie la « **chaîne du livre numérique** ».

Les principaux **processus métiers** à l'œuvre dans la production et la distribution de livres numériques peuvent être définis par les six catégories suivantes :

1. **Production** des livres numériques, qui rassemble les fonctions traditionnelles d'écriture et d'édition des livres et les fonctions techniques de fabrication d'une version numérique à des formats de fichiers variés.
2. **Archivage** des fichiers d'ebooks et ajout des métadonnées associées permettant de décrire et d'indexer les livres. Les fichiers et métadonnées, archivés durablement et de manière organisée sur des serveurs ad-hoc, doivent pouvoir être retrouvés à tout moment par l'éditeur en vue d'utilisations diverses pour lui-même ou pour tous ses partenaires de la distribution numérique.
3. **Catalogage et Recherche** des ebooks et des métadonnées sur un site en offrant une recherche dans une base de livres, grâce notamment aux métadonnées et à la recherche plein texte permise par la numérisation. Ce catalogage permet d'une manière générale à des utilisateurs tiers (internauts, distributeurs professionnels) de trouver les sélections de livres correspondant à leurs critères de recherche et de consulter éventuellement des pages ou des extraits limités de livres correspondant à cette recherche.
4. **Diffusion** des ebooks (texte intégral) avec des technologies de protection des droits d'auteur, des solutions de lecture et des modèles commerciaux variés. Le numérique introduit une étroite corrélation entre ces trois points : un modèle commercial y est toujours lié à des solutions de lecture spécifiques et est garanti par des mesures de protection. D'où le rôle d'acteurs nouveaux spécialisés dans la fourniture de ces solutions économique-techniques intégrées : les diffuseurs numériques, ou « agrégateurs ».
5. **Vente** des livres numériques, qui correspond au rôle tout à fait traditionnel des librairies, incluant la prise en charge de la commande, du paiement, de la livraison et du service après-vente.
6. **Prêt** des livres numériques, qui correspond au rôle traditionnel des bibliothèques, incluant notamment la mise à disposition de collections de livres numériques de manière gratuite ou avec un abonnement forfaitaire à une collectivité identifiée de lecteurs.

Cette description théorique présente des processus métiers ou rôles fonctionnels, qui peuvent être assurés par des acteurs spécialisés, sachant qu'un même acteur peut assurer plusieurs de ces fonctions, ou se regrouper avec d'autres pour les assurer, et que plusieurs types d'acteurs peuvent accomplir une même fonction. La description des différents

processus métiers peut être complétée par une typologie des principaux acteurs, de leur spécialisation et de leurs relations au sein de la chaîne du livre numérique :



Ce schéma n'est pas le seul « arrangement » possible des relations entre différents acteurs et des regroupements possibles de fonctions pour un seul acteur. Mais il semble être l'un des schémas optimaux dans l'organisation en cours du secteur. Il servira de point de repère pour définir le rôle que devrait jouer *Europeana* au sein de cette chaîne pour les livres sous droits.

Le schéma tel quel semble assez bien correspondre à une partie du circuit de l'édition numérique aux États-Unis, où des acteurs spécialisés uniquement dans la diffusion numérique ou dans l'archivage de données sont apparus, de même que de purs « ebook stores » ou librairies de livres numériques proposées par des librairies traditionnelles. Toutefois, plusieurs agrégateurs y jouent également le rôle de librairie numérique et bien entendu, gèrent de ce fait une sorte de portail de recherche. En Allemagne, les éditeurs et les libraires regroupés ont créé en 2007 une plate-forme commune en principe destinée au seul archivage des fichiers et métadonnées, « VolltextSuche Online », qui a vocation à négocier collectivement des accords avec plusieurs spécialistes de diffusion et des portails de recherche, bien qu'il semble qu'elle se soit orientée pour le moment vers un accord avec un diffuseur unique. En France, la création d'une plate-forme d'archivage commune des diffuseurs est un projet en discussion mais non encore initié, sachant que des plates-formes pourraient également s'organiser autour de plusieurs groupes d'éditeurs ou d'un diffuseur. Partout, les accords ont lieu avec quelques diffuseurs numériques spécialisés, qui à leur tour développent des accords commerciaux avec des bibliothèques numériques pour le prêt

et, éventuellement, de librairies numériques tierces. La fonction particulière de catalogage et recherche, largement partagée par les autres acteurs numériques pour leur propre catalogue, a fait l'objet ces dernières années d'initiatives de purs acteurs de l'Internet ayant déjà des positions dominantes en tant que moteurs de recherche.

Ces fonctions et acteurs typiquement numériques doivent être complétés par les deux suivants, qui réintroduisent le papier dans la chaîne, pour représenter complètement les usages des textes numérisés :

- les imprimeurs, qui de plus en plus utilisent comme fichiers source des fichiers de type PDF de même nature que ceux qui sont utilisés pour les ebooks. L'impression à la demande (à l'unité) est en particulier une forme d'exploitation très intéressante des livres numériques envisageable à l'avenir ;
- les librairies traditionnelles, qui peuvent bénéficier des indications fournies aux internautes par les portails de recherche renvoyant vers ces librairies pour acheter la version imprimée.

2.3 Typologie des modèles économiques

2.3.1 La commercialisation des contenus numériques sous droits

L'idée que les livres, comme tous les biens culturels, une fois numérisés, devraient ou pourraient être distribués gratuitement, est parfois avancée avec notamment les arguments hétérogènes suivants, repris, en les combinant de manière plus ou moins cohérente, par des acteurs aux intérêts souvent objectivement divergents :

- a. La numérisation des livres les transformerait en « biens non rivaux », dont la consommation par un individu ne nuit pas à celle d'un autre, ou en « biens d'information », dont la tarification au coût marginal, en accord avec la théorie micro-économique néoclassique dans un cadre de concurrence « pure et parfaite », conduirait théoriquement à un prix de marché nul. Les coûts de reproduction et de distribution des livres numérisés sont en effet dans ce cas présentés comme quasiment nuls.
- b. Internet serait un univers où la gratuité est naturelle et représenterait une forme nouvelle d'économie, l'économie de l'accès (selon le titre de l'ouvrage de Jeremy Rifkins, *L'âge de l'accès*, La découverte, 2000), où les internautes-consommateurs ne payent plus que pour des services (l'accès au web) et non pour des produits (les livres) et où les internautes-citoyens bénéficient d'un accès sans limite à l'information et à la culture.
- c. La diffusion gratuite des livres numériques serait un outil de promotion qui faciliterait en fait la commercialisation de leurs versions physiques.
- d. Le modèle économique du financement des biens culturels numériques devrait être le modèle publicitaire adopté par les médias.

La controverse, très nourrie dans la presse et sur Internet, sur les relations entre numérique et gratuité, dépasse largement le cadre de cette étude aux visées pratiques. On peut cependant remarquer que de solides contre-arguments économiques s'opposent à chacun des arguments présentés ci-dessous, dont certains sont développés par exemple dans deux essais récents : *Gratuit ! Du déploiement de l'économie numérique* d'Olivier Bomsel (Gallimard, 2007), *La gratuité, c'est le vol* de Denis Olivennes (Grasset, 2007) :

En reprenant point par point :

- a. Le modèle néoclassique de la tarification au coût marginal n'est pas adapté au cas des biens immatériels dont les fonctions de production présentent des rendements croissants : plus on en distribue, plus les coûts unitaires diminuent. Le numérique ne change ici rien de fondamental par rapport au monde physique : si un livre imprimé était vendu à son coût marginal, il faudrait se contenter de payer son impression et sa distribution. Un autre exemple est celui des inventions, heureusement protégées par des brevets pour favoriser l'innovation. Les coûts fixes de création et d'édition d'une œuvre numérisée méritent bien sûr une juste rémunération, sans laquelle les créations intellectuelles et le travail éditorial ne pourraient plus exister dans une économie de marché. Et leur prix, régulé par des mécanismes concurrentiels, reflète cette rémunération, dont la légitimité et la nécessité économique ne disparaissent pas avec la numérisation. Sans compter que les coûts marginaux de conservation et de distribution ne sont jamais réellement nuls. Enfin, la valeur d'un bien n'est pas liée seulement à ses coûts de production mais aussi à sa valeur d'usage, qui peut être grande même lorsque ses coûts marginaux de production sont nuls. Toute l'économie des biens culturels est fondée depuis le XIX^e Siècle sur la rémunération des auteurs et sur le rôle des éditeurs, et ce modèle économique a permis l'essor formidable des industries culturelles depuis cette époque ainsi que l'indépendance des auteurs de la création intellectuelle et artistique par rapport aux mécènes ou à l'Etat.
- b. L'idée qu'Internet serait un univers de gratuité naturelle ou que l'économie de l'accès y ferait disparaître la vente unitaire de produits est une idée naïve si on veut en faire une idée générale. Cette idée a surtout caractérisé les débuts de l'Internet, où régnait une certaine forme de confusion économique sous l'influence des premiers aficionados qui y voyaient l'émergence d'un monde nouveau où régnerait désormais l'abondance, sans s'interroger sur les conditions de raccordement de cette économie de l'abondance avec l'autre économie, toujours dominante, des biens payants et des revenus nécessaires pour y survivre. Internet augmente certainement l'offre de ressources culturelles gratuites, en particulier celles des moteurs de recherche, des encyclopédies collaboratives et de multiples sites d'information. Mais la gratuité est également bien souvent un « cheval de Troies » permettant à des industries de télécommunication ou à divers types d'acteur de l'Internet de déployer avec plus d'efficacité les effets de réseaux propre au numérique... et d'en tirer des bénéfices bien plus importants à d'autres endroits de la chaîne. Surtout, Internet n'est qu'un nouvel outil de communication et de distribution, largement neutre par lui-même. Rien ne le prédestine à favoriser plutôt la distribution de services gratuits que de produits payants, ce qu'il fait très bien pour les voyages, les ordinateurs ou les livres imprimés.

- c. L'argument de la promotion des livres papier par leur version numérique suppose implicitement que la lecture numérique n'a aucune valeur par elle-même en tant que mode de lecture final des œuvres : sous un jour « moderniste », elle va en fait à l'encontre de l'idée, bien plus moderne et réaliste à l'avenir, que la lecture numérique est un authentique mode de lecture alternatif qui pourra de plus en plus souvent dispenser de la lecture de l'œuvre sur un support physique ; surtout, et ce qui nous importe pour la présente étude, cet argument est jugé fantaisiste par la plupart des éditeurs.
- d. Le financement des livres numériques par la publicité est une idée bien plus intéressante et raisonnable, qui donne déjà lieu à certaines expériences qui devraient se développer dans l'avenir. Mais les ressources publicitaires disponibles sur Internet ne sont pas extensibles et celles qui seraient disponibles pour financer les livres en ligne seraient certainement, en tout cas pour le moment, très insuffisantes : elles ne le sont même pas pour la presse. D'autre part, ce mode de financement ne semble pour le moment adapté qu'à certaines catégories d'ouvrages (le pratique ou les voyages par exemple). Ni les éditeurs ni les lecteurs ne sont prêts à en accepter l'extension à l'ensemble des livres.

Il n'en demeure pas moins que **la dématérialisation des contenus enrichit considérablement les possibilités de modèles économiques pour tarifer les biens culturels** : livres, musique, films, photographies. Elle permet aussi bien, et c'est encore ce qui est largement souhaité, de vendre les livres à l'unité selon le modèle de l'exemplaire, avec des prix toutefois réduits et des granularités plus fines (chapitres, morceaux de musique), que de s'approcher de l'« économie de l'accès » en proposant des abonnements à des « bundles » vendus avec des appareils de lecture ou à des corpus importants de textes accessibles de manière illimitée. Elle permet tout autant, et c'est une bonne chose pour leur rôle de médiation sociale et culturelle, de préserver la stricte médiation économique des bibliothèques traditionnelles, que d'imaginer des modèles économiques conduisant à un plus grand continuum entre bibliothèque et librairie. Nous présenterons dans l'étude un inventaire de la diversité des modèles économiques rendus possibles par le numérique.

Si on considère la seule présentation d'extraits de documents numériques à **des fins de recherche et de consultation partielle**, soit la fonction de « catalogue en ligne », deux modèles économiques simples sont envisageables :

- un **modèle gratuit**, dans lequel les usagers ont la possibilité de faire des recherches dans la base de textes et d'afficher une ou plusieurs pages du livre trouvé, sans rémunération de la plate-forme ni des ayants droit, qui peuvent y trouver un avantage consistant dans la promotion de leurs ouvrages ;

- un **modèle payant**, dans lequel les usagers bénéficient des mêmes possibilités, avec peut-être des fonctionnalités étendues (impression des pages trouvées, récupération des métadonnées sous la forme de fichiers), mais avec un paiement de la plate-forme et/ou des ayants droit pour la fourniture de cette information. Ce paiement peut être forfaitaire ou lié à l'usage effectif (nombre de pages consultées ou imprimées).

Le modèle gratuit prévaut aujourd'hui largement, en particulier pour les catalogues, portails et moteurs de recherche en ligne destinés au grand public. Le modèle payant est toutefois pratiqué par des bases de données bibliographiques destinées aux professionnels, comme la base *Electre*, ou aux chercheurs dans des domaines spécialisés.

Si on considère par contre la distribution de documents numériques **en version texte intégral**, la taxonomie des modèles économiques possibles est plus complexe. On peut d'abord distinguer deux grandes catégories :

- les **modèles de librairie numérique**, ou de **licences individuelles**, c'est-à-dire à destination directe du public qui paye pour les contenus achetés, par l'intermédiaire typiquement d'une librairie ;

- les **modèles de bibliothèques numériques**, ou de **licences groupes**, c'est-à-dire à destination d'entreprises ou d'organismes, typiquement des bibliothèques, qui proposent à un groupe d'individus un accès mutualisé à des documents numériques diffusés par un offreur. En principe, dans ce modèle, les usagers ne paient pas eux-mêmes pour l'accès aux contenus, sauf éventuellement sur une base forfaitaire déconnectée des coûts directs de ces contenus. Les bibliothèques effectuent ainsi une médiation économique (et culturelle) entre les offreurs (éditeurs - agrégateurs - libraires) et les lecteurs, en transformant des biens privés en ressources collectives. Le prêt est l'outil traditionnel de cette médiation.

2.3.2 Les modèles économiques de librairie numérique

On peut croiser cette distinction avec les distinctions suivantes, concernant les droits des usagers finaux :

- l'attribution de **droits d'usage pérennes** sur les documents numériques, qui consiste en une vente au sens strict habituel ;
- l'attribution de **droits d'usage temporaires** sur les documents numériques, qui représente économiquement un « prêt » si elle est gratuite, ou une « location » si elle est payante, même s'il n'y a pas de transfert de fichiers. Ce prêt peut être « matérialisé » par le transfert par téléchargement d'un fichier « chrono dégradable » ou par la limitation de la consultation en ligne en termes de « sessions de lecture » et de durée d'appartenance au groupe. **Il ne faut pas confondre le mode de lecture, hors ligne ou en ligne, avec ou sans téléchargement, avec la durée des droits attribués.**

La vente peut s'effectuer sous deux formes en ce qui concerne les contenus :

- la **vente titre par titre**, permettant aux clients de choisir les livres qu'ils achètent un par un comme dans une librairie habituelle ;
- la **vente en collections ou « packages »**, par exemple l'ensemble des ouvrages de tel éditeur important pour la médecine ou l'informatique, ou toute la base de données d'un agrégateur dans le domaine du pratique.

Enfin, les droits peuvent porter sur :

- des **livres complets**
- des **parties de livres** : chapitres, articles

On aboutit donc à 8 catégories théoriques de modèles économiques pour les licences individuelles, présentées dans le tableau suivant :

Type de modèle	Droits pérennes		Droits temporaires	
	Titre par titre	Collections	Titre par titre	Collections
Livres complets	L1	L2	L3	L4
Chapitres, parties	L5	L6	L7	L8

Parmi ces 8 catégories théoriques, seules L1, L3, L4, L5, L7 sont en fait pertinentes et proposées en pratique. L1 correspond à la reproduction du modèle classique de vente à l'exemplaire. Dans les cas L3 et L7, qui peuvent en fait se confondre, les durées des droits sont en général courtes : il s'agit de proposer des solutions dites de « pay-per-view ». Dans le cas L4, elles sont plutôt longues : il s'agit de proposer des accès forfaitaires à de larges collections de livres pendant par exemple un mois ou un an.

Certaines librairies de livres numériques proposent également des modèles mixtes entre L1 et L2 en offrant des « bundles » d'ebooks rassemblant par exemple les titres d'un auteur.



Newsletter du site eReader.com

2.3.3 Les modèles économiques de bibliothèque numérique

Concernant les modèles de licences groupe proposées aux bibliothèques, on peut faire les distinctions suivantes :

- En termes de pérennité des droits de prêt ou des licences groupe elles-mêmes :
 - les bibliothèques peuvent **acheter définitivement** les livres numériques ;
 - elles peuvent souscrire un **abonnement** leur permettant de les prêter pendant une période.

- En termes de contrôle des accès aux livres, on peut distinguer :
 - les modèles dits à **accès simultanés limités**, déterminant un nombre fixé par contrat d'accès simultanés aux livres, jouant le même rôle économique et fonctionnel que les « exemplaires » dans le domaine physique ;
 - les modèles dits à **accès simultanés illimités**, dans lequel les accès sont libres pour une communauté d'utilisateurs et pendant la période de la licence collective. Les accès ne sont pas illimités stricto sensu, mais dans le cadre d'une population d'usagers et d'une période : tous les utilisateurs identifiés peuvent se connecter simultanément sur les ressources tant qu'ils sont membres de la collectivité acheteuse et tant que dure l'abonnement.

- En termes de contenus, on peut à nouveau distinguer :
 - la **vente titre par titre**, permettant aux bibliothèques de constituer leur collection de livres numériques en les sélectionnant en fonction de leurs choix documentaires ;
 - la **vente en collections ou « packages »**.

On peut proposer le tableau de synthèse suivant des 8 principaux types de licences groupes :

Type de modèle	Achat pérenne		Location annuelle	
	Titre par titre	Collections	Titre par titre	Collections
Accès simultanés limités	B1	B2	B3	B4
Accès simultanés illimités	B5	B6	B7	B8

Parmi ces 8 catégories théoriques, B1, B3, B4, B8 sont en fait les plus régulièrement proposées en pratique. B1 correspond à la reproduction du modèle classique de vente d'exemplaires que la bibliothèque pourra prêter proportionnellement à ce nombre d'exemplaires. B3 est la déclinaison en abonnement de ce modèle traditionnel. B4 permet de proposer à la bibliothèque l'équivalent de « jetons de prêt » interchangeables entre des livres d'une collection. B8 est le modèle répandu notamment pour les revues scientifiques, et propose un accès illimité à une base de livres d'un thème donné, consultables librement

par une population d'utilisateurs définis (par exemple les étudiants d'une université) pendant une certaine période (par exemple un an).

Remarques sur les modèles de bibliothèque numérique :

La distinction entre la vente de livres complets ou la vente de parties, bien que possible, n'est en général pas pertinente pour définir des modèles de bibliothèque numérique : dans la plupart des cas, même si le niveau des chapitres est considéré pour le prêt, ou pour la création à la demande de dossiers rassemblant des parties de livres, la bibliothèque elle-même achète ou s'abonne à des livres entiers.

La distinction entre l'achat pérenne et l'abonnement annuel de la part des bibliothèques ne doit pas être confondue avec la distinction entre les droits pérennes et les droits temporaires achetés par des individus. Le modèle de bibliothèque numérique est en effet un modèle de type « B to B to C », qui ajoute un niveau de plus au modèle de librairie numérique. Au premier niveau (relation entre les bibliothèques et les usagers), les modèles de bibliothèque sont toujours des modèles de droits d'usage temporaire pour les usagers, au minimum limité à la durée d'appartenance à la communauté d'usagers. C'est au deuxième niveau (relation entre le fournisseur et les bibliothèques) qu'est introduite ici une distinction entre achat et location.

Tous les modèles de bibliothèques numériques sont, du point de vue de la théorie économique, des modèles de « prêt numérique », puisque les droits conférés aux usagers sont par nature temporaires et que cet accès temporaire est gratuit pour l'utilisateur. La nature de cette durée peut être variable : elle peut reproduire la durée fixe d'un prêt habituel en bibliothèque, par exemple une semaine, ou reproduire la durée variable d'une consultation sur place, par exemple, le temps qu'un lecteur choisit pour la lecture, borné éventuellement par des sessions de lecture maximales de 24 heures.

Il est fréquent que cette distinction entre **prêt** à un usager et **achat** soit confondue, à tort, avec deux autres distinctions que nous avons rencontrées :

- la distinction entre **téléchargement** et **lecture en ligne** ;
- la distinction entre **accès sur place** dans les locaux d'une bibliothèque et **accès à distance** sur un site Internet.

Cette confusion est liée d'une part aux aspects symboliques du téléchargement (qui ressemble à un prêt d'objet, les fichiers), et d'autre part à l'opposition prêt / consultation dans le langage des bibliothécaires. Elle conduit à de sérieuses erreurs quand il s'agit d'analyser la pertinence et la signification des différents modèles économiques envisagés. En fait, les 8 modèles économiques présentés peuvent tous être croisés fonctionnellement avec les deux distinctions techniques que nous venons de citer et sont donc totalement indépendants d'elles.

Il est vrai toutefois que les modèles à accès simultanés limités impliquent en général des restrictions de durée de prêt plus fortes que les modèles à accès simultanés illimités, puisqu'ils impliquent de gérer une plus grande rareté des ressources collectives. Les modèles à accès simultanés illimités n'impliquent qu'une restriction liée à la durée des abonnements collectifs et à la durée d'appartenance à la collectivité des lecteurs autorisés. La notion de

« prêt numérique stricto sensu » pourrait alors être réservée aux modèles à accès simultanés limités B1 à B4. L'idée que ces modèles de prêt, entendu ainsi, seraient par nature moins adaptés aux contenus numériques que les modèles à accès simultanés illimités, est parfois reprise avec l'argument brut suivant : « à quoi sert de proposer des contenus numériques si c'est pour reproduire les modèles de l'imprimé ? ». Cet argument est toutefois très peu pertinent ; il reproduit à un autre niveau l'erreur qui assimile les biens d'information à des biens gratuits : certes le numérique permet, techniquement, de gérer de multiples accès simultanés avec peu de surcoûts physiques (quoique cela soit nettement moins vrai s'il s'agit d'offrir des accès en ligne à des vidéos par exemple). Mais le tarif des ressources numériques ne reflète pas leur rareté physique, mais l'équilibre entre la demande de lecture et l'offre de détenteurs de droits sur des œuvres intellectuelles. La bonne question est par conséquent : **quel est le meilleur mode de calcul pour définir un tarif d'équilibre ?**

La réponse des modèles à accès limités est de tenir compte « au plus près » de la demande, en adaptant le tarif à l'intensité de la consommation, mesurée par les accès simultanés autorisés. La référence au modèle papier offre par ailleurs un point de repère utile pour définir le « juste prix ». La réponse des modèles à accès illimités est de mesurer cette intensité de consommation par une variable qui la mesure de manière plus approximative : le nombre d'utilisateurs. Les modèles à accès illimités fixent en fait le tarif d'équilibre en décidant d'ignorer la valeur de la consommation moyenne par usager à un moment donné. Il en résulte une incertitude plus grande sur la consommation réelle offerte par le fournisseur et donc un modèle de marché faisant une part plus grande au pouvoir de négociation respectifs des offreurs et des demandeurs pour fixer le prix. De là découlent aussi souvent ce qui a été appelé les « big deals », c'est-à-dire les marchés à prix très élevés et inflationnistes imposés par le pouvoir de négociation de certains éditeurs quasi-monopolistiques. Comme quoi la rhétorique de la non-rareté du numérique peut en fait se retourner contre les usagers. Ce qui ne signifie pas qu'il n'est pas possible de proposer des « deals » d'accès illimités équilibrés, mais que cet équilibre est par nature plus difficile à définir.

2.3.4 Vers des modèles hybrides ?

L'irruption des nouvelles technologies dans le monde du livre et des bibliothèques implique que ces dernières acceptent de remettre au moins partiellement en question certaines de leurs pratiques ou de leurs missions traditionnelles, pour s'adapter à de nouvelles logiques et à de nouvelles demandes. Le concept de « bibliothèque hybride » est ainsi apparu au cours des dernières années, avec le développement dans les bibliothèques des services de documentation et d'aide à la recherche fondés sur les technologies numériques et plus encore des bibliothèques numériques proprement dites. La bibliothèque, qui conserve son rôle essentiel de médiation entre l'offre et la demande de lecture, voit ses métiers évoluer lorsqu'elle mêle les ressources documentaires traditionnelles à ces nouvelles ressources numériques. En ce qui concerne les documents numériques en effet, son rôle de négociateur de tarifs et de sélection des offres électroniques devient de plus en plus important, tandis que ses missions de sélection et surtout de conservation des documents le sont sans doute moins.

Mais un autre type d'hybridation possible est moins souvent commenté. Il concerne les modèles économiques séparant traditionnellement une librairie d'une bibliothèque. Un

dernier modèle, encore peu fréquent, peut ainsi être ajouté pour compléter la typologie des modèles économiques possibles de bibliothèque numérique : **la vente de droits de consultation (B9)**, mesurés par un nombre de téléchargements (pour des durées de lecture prédéfinies), ou par un nombre de pages consultées ou imprimées.

Ce modèle va plus loin que les modèles à accès simultanés limités en ce qui concerne la mesure de la consommation effective des documents numériques, puisqu'il repose sur une tarification liée à la mesure directe de cette consommation. Dans les modèles à accès limités, deux paramètres de consommation demeurent incertains : la durée de prêt, choisie par la bibliothèque, qui détermine le nombre total de prêts possibles pour un même nombre d'accès simultanés sur une période, et le nombre de prêts simultanés effectivement réalisés, qui peut être inférieur au nombre maximal autorisé. Dans les modèles de ventes de droits de consultation, au contraire, chaque prêt, d'une durée donnée ou d'un nombre de pages fixé, donne lieu à un paiement convenu. Ce modèle présente des avantages pour les bibliothèques et pour les éditeurs : il limite pour les bibliothèques le risque de payer pour des ressources numériques sous-utilisées, risque qui est maximal dans les modèles à accès illimités ; il garantit réciproquement aux éditeurs que leur rémunération dépendra directement de l'usage de leurs livres, même si la population d'utilisateurs est très large, limitant ainsi les risques économiques de la « mutualisation » des accès si plusieurs bibliothèques s'associent pour créer une bibliothèque numérique commune.

Ce faisant, ce modèle conduit à une évolution dans la nature du rôle économique des bibliothèques : il se rapproche des modèles de librairie, en particulier du modèle L3. **La bibliothèque joue en effet dans ce cas un rôle de tiers payant se substituant aux lecteurs dans le modèle de pay-per-view, un peu comme la sécurité sociale ou une mutuelle se substitue aux patients pour le règlement des frais médicaux.** C'est donc un modèle plus éloigné que les autres de la pratique habituelle d'une bibliothèque, dont l'avenir dira s'il est retenu parmi les modèles économiques nouveaux du numérique. A titre d'exemple, c'est le modèle que propose la chaîne ARTE pour vendre des vidéos numériques aux bibliothèques.

Enfin, une autre forme d'hybridation est encore plus nette et proposée par certains fournisseurs : **la combinaison entre un modèle de bibliothèque**, en général de type B8, pour la lecture en ligne des textes, **et un modèle de librairie**, de type L3 ou L7, pour l'impression de pages ou le copier-coller. Les lecteurs bénéficient alors d'un accès illimité aux contenus auxquels la bibliothèque s'abonne, mais doivent payer pour chaque impression ou copier-coller de pages sur le mode du pay-per-view. Le paiement s'effectue directement dans ce cas auprès du fournisseur, qui en quelque sorte mêle sa librairie à la bibliothèque numérique de son client.

2.4 Les mesures techniques de protection des droits d'auteur ou « DRM »

2.4.1 L'utilité des DRM ?

Autre sujet de débats multiples et de controverses, ayant entouré notamment en France le vote en 2006 de la loi « DADVSI », les mesures techniques de protection des fichiers numériques ou « DRM » pour Digital Rights Management sont, d'une manière assez consensuelle dans le milieu éditorial, le garant nécessaire de la mise en œuvre des modèles économiques présentés ; en tout cas tant que d'autres formules satisfaisantes de protection des droits d'auteur n'auront pas été trouvées et fait leurs preuves auprès des ayants droit.

Les DRM sont pour l'essentiel destinés à empêcher une dissémination incontrôlée des œuvres sur les réseaux peer-to-peer, et non directement à empêcher la copie privée. Le numérique introduit en effet une rupture technologique radicale dont on ne peut pas ne pas tenir compte : **la possibilité de reproduire instantanément le support d'une œuvre et d'en diffuser des millions de copies en très peu de temps à coût quasi-nul**. Ignorer ce fait au nom de la liberté du consommateur et de la défense d'une exception au droit d'auteur née à une époque où cette technologie n'existait pas serait très peu raisonnable. Car le changement d'échelle de la copie introduit un vrai changement de nature, qui est susceptible de détruire les industries culturelles ou du moins de bouleverser totalement leur organisation, et notamment celle de la chaîne du livre. D'un point de vue strictement économique, la double question suivante résume ce risque concret :

- est-il concevable que des individus continuent à acheter des livres s'ils peuvent se les fournir gratuitement grâce à un système de copies illimitées ?
- comment les ayants droit et professionnels de la chaîne du livre peuvent-ils être rémunérés si le marché des livres est peu à peu cannibalisé par des offres gratuites de copies ?

La réponse à cette double question, qui comporte des aspects techniques, économiques et politiques, sort à nouveau largement du cadre de cette étude, mais les remarques pratiques qui suivent justifient, si cela était nécessaire, le parti pris résolu en faveur de la nécessité des DRM pour la diffusion des livres numériques sous droits au stade actuel du marché, du droit et des schémas techniques disponibles :

- aucun modèle économique satisfaisant de vente d'ouvrages numériques sans DRM n'est concrètement apparu jusqu'à présent, hormis quelques expériences de financement publicitaire des livres, dont les limites ont déjà été soulignées.
- la directive européenne sur le droit d'auteur dans la société de l'information incite les Etats de l'Union Européenne à fournir un cadre légal aux mesures de DRM et aux sanctions contre toute tentative de détourner ces mesures ; la loi DADVSI a transposé cette directive en légalisant et en encadrant l'application de ces mesures, tout en prévoyant certaines exceptions.
- A la différence d'expériences récentes qui semblent s'esquisser sur le marché de la musique, très différent dans son historique et ses problématiques, quasiment aucun éditeur de livres ne s'engagerait aujourd'hui dans la diffusion de livres numériques de manière importante sans aucun DRM ou mesure de protection crédible, même s'il est vrai que leur

sensibilité à cette question et la rigueur des mesures exigées présentent des degrés variables. Les rares exceptions ne permettraient pas de créer une offre significative.

Un autre objectif de la protection technique des contenus numériques, auquel les auteurs sont particulièrement sensibles, est par ailleurs lié au **droit moral sur les œuvres** : s'il est possible très facilement et sans limite de copier-coller des parties d'ouvrages pour les éditer dans un traitement de texte en les incorporant à ses propres documents, le numérique risque de devenir un outil de plagiat à grande échelle ou de déformation de textes originaux, qui va bien au-delà du droit de citation inscrit dans le code de la propriété intellectuelle. Les auteurs et les éditeurs souhaitent en général que des mesures techniques limitent ce risque, potentiellement accru par le numérique, sans évidemment pouvoir l'empêcher plus qu'avec des ouvrages imprimés.

Bien entendu, tout système de DRM est un jour ou l'autre « craquable », plus ou moins facilement selon la technologie et le niveau de cryptage adoptés. Ce n'est pas une raison suffisante pour rejeter leur usage. D'une part à la loi « tout DRM peut être craqué » a toujours répondu jusqu'à présent la loi « tout crac peut être dépassé par un DRM mis à jour », du moins lorsque l'éditeur des DRM avait un intérêt suffisant à le faire compte tenu du marché. Il s'agit donc plutôt d'un mouvement continu, connu d'ailleurs des cryptologues depuis des siècles, entre les codeurs et les décodeurs, dont l'issue n'est jamais définitive ni dans un sens ni dans l'autre. D'autre part, et c'est en fait leur plus modeste et principale fonction, le niveau de sécurité atteint par les DRM est en général suffisant pour dissuader la plupart des utilisateurs, même si un crac circule – temporairement si la loi est appliquée – parmi les plus experts d'entre eux.

Il est bien entendu également souhaitable que les DRM ne s'opposent pas à un usage souple et commode des livres numériques dans des environnements technologiques variés. A ce titre, les DRM actuellement utilisés pour le téléchargement restreignent la possibilité de copie privée mais ne l'empêchent pas : ils la limitent à un certain nombre de copies, c'est-à-dire d'installations du fichier du livre numérique sur différentes machines « activées » avec un même compte utilisateur. Certains systèmes permettent de désactiver le livre sur une machine pour l'installer sur une autre, assurant ainsi la possibilité de faire circuler le livre numérique comme un objet physique.

2.4.2 Les DRM pour la lecture en ligne

La lecture de contenus numériques en ligne fait souvent l'objet de mesures de protection bien plus faibles que celles associées aux contenus téléchargés, voire inexistantes.

Une première raison peut être liée à la volonté délibérée de ne pas protéger ces contenus et au contraire d'offrir les fonctionnalités d'usage les plus larges possibles aux usagers, comme le propose par exemple *Europeana* sur son site avec le Reader utilisé pour les livres du domaine public. Ce Reader propose de nombreuses fonctions très commodes : impression, copier-coller, sauvegarde du document, téléchargement, mémorisation de l'adresse URL fixe des pages du document. Néanmoins toutes ces fonctions permettent une copie et une diffusion du document sans protection.

Une deuxième raison peut être liée à l'idée, vraie, que la lecture en ligne offre une certaine protection naturelle plus forte que le téléchargement car les pages sont lues une par une

sans que le fichier de l'ouvrage complet ne soit jamais transmis (sauf si la fonction « téléchargement » est ajoutée). C'est aussi l'une des raisons pour laquelle les fichiers musicaux ou les vidéos numériques sont parfois proposés en « streaming ». Néanmoins, cette protection naturelle n'est pas suffisante pour garantir aux éditeurs une sécurité crédible quant aux usages de leurs publications numériques en dehors des modèles économiques acceptés. En effet, il est assez facile, sans DRM, de reconstituer tous les livres en ligne d'une collection en enregistrant les pages une ou à une à l'aide de systèmes automatisés tels que des robots qui simulent l'activité d'un lecteur et enregistrent toutes les pages.

Il est donc souhaitable, et possible, d'ajouter des DRM aux logiciels de lecture en ligne destinés à la lecture de livres numériques sous droits. Les mesures de protection utilisables sont en fait assez nombreuses. On peut par exemple ajouter les types de sécurité suivants :

- Donner accès à la lecture de fichiers images plutôt que de fichiers texte (mis en page). En effet :
 - d'une part les fichiers images, s'ils sont récupérés, ne sont pas directement éditables par les usagers dans un logiciel de traitement de texte, sauf à procéder à une reconnaissance optique page par page et à une correction des erreurs, ce qui en accroît fortement la difficulté.
 - d'autre part, un fichier texte mis en page permet une reproduction totalement fidèle du livre original avec potentiellement une qualité d'impression parfaite ; un fichier image au contraire, surtout si on se limite à une résolution de 72dpi suffisante pour une bonne lecture à l'écran, ne permet pas une reproduction de même qualité pour celui qui voudrait l'industrialiser.
 - enfin, on peut mieux protéger les fichiers image que les fichiers texte, qui sont bien plus faciles à intercepter dans les flux réseaux.

Cette mesure semble aller à l'inverse de la démarche accomplie par la BnF en proposant désormais des fichiers auxquels a été appliquée une reconnaissance optique de caractères (« OCR » en anglais) afin de permettre l'indexation et la recherche plein texte. En fait, proposer la lecture en mode image n'empêche nullement de permettre une recherche plein texte, qui sera effectuée par ailleurs dans un fichier texte non affiché à l'utilisateur, avec les résultats correspondants mis en surbrillance aux endroits appropriés de l'image.

- Rendre extrêmement difficile pour un usager normal l'enregistrement des pages affichées en empêchant non seulement l'usage de la fonction « enregistrer sous » mais aussi la récupération du contenu de la page ou des éléments externes appelés par la page au sein des sources HTML que le navigateur peut afficher.
- Contrôler les droits de copier-coller, en limitant par exemple le nombre de copier-coller possibles, par jour ou session de lecture d'un usager.
- Contrôler de la même manière les droits d'impression.
- Contrôler la durée de lecture, afin de définir des sessions de lecture maximale, et de crypter la page de manière différente régulièrement.

- Garantir un accès sécurisé aux pages lues en ligne, sous la forme d'une procédure d'authentification du lecteur afin de contrôler ses droits de lecture et d'éviter que les pages ne soient accessibles à tout internaute.
- Enfin, utiliser des systèmes anti-robots du type tests « captcha » (*Completely Automated Public Turing test to Tell Computers and Humans Apart*), ou analyse du rythme de lecture afin de détecter et de bloquer la lecture automatisée par un robot visant à récupérer les pages une à une pour reconstituer les livres.



Exemple de texte « captcha »

A noter : certains types de captcha sont définis sans image, par exemple avec des expressions à compléter, de manière à satisfaire des normes d'accessibilité pour les handicapés visuels.

On peut remarquer que ces mesures, si elles permettent un meilleur contrôle de l'usage des livres numériques en ligne, restreignent en fait assez peu l'usage normal et confortable des lecteurs. Les environnements technologiques compatibles pour la lecture en ligne sont très variés. **La possibilité de lecture nomade de son livre, en changeant d'ordinateur mais en s'authentifiant avec le même compte, est préservée.** La notion de copie numérique privée n'a par ailleurs guère de sens dans ce contexte où aucun fichier n'est téléchargé, sauf à l'associer à la possibilité pour un client authentifié de « prêter » son authentification à un autre pour le temps d'une lecture.

2.4.3 Les DRM pour le téléchargement

Bien que le téléchargement conduise à récupérer sur une machine un fichier complet, la protection des fichiers par DRM présente en général un caractère robuste, lié à des algorithmes de cryptage plus sophistiqués que ceux utilisés par la plupart des logiciels de lecture en ligne. Chaque format de lecture est associé à une technologie DRM différente, qui est en général compatible avec différents types de machines ou de systèmes d'exploitation. Par exemple le système de DRM de Mobipocket, qui permet de protéger les fichiers au format PRC, est compatible avec des appareils de lecture très variés, tels qu'ordinateurs, PDA de type Pocket PC ou Palm, Smartphones équipés du système Symbian, et très prochainement l'Illiad™ de Irex, appareil de lecture à base d'encre électronique. Les livres numériques au format PRC peuvent circuler librement entre plusieurs de ces appareils où est installé le logiciel Mobipocket Reader activé avec le même compte client. D'autres systèmes DRM, comme celui par exemple utilisé par Sony pour son appareil de lecture à base d'encre électronique, ne sont compatibles qu'avec un seul type de machine qui ne peut lire des fichiers protégés qu'avec ce type de DRM, pour

des raisons stratégiques de captation de marché, sur le mode de la stratégie iPod™/iTune (stratégie dite du « winner takes all »). Cette stratégie est toutefois en contradiction d'une part avec l'intérêt des éditeurs, d'autre part avec les exigences d'interopérabilité contenues dans les dispositions légales récentes de plusieurs pays européens, et ne durera donc peut-être pas.

Les principaux types de sécurité qui peuvent être ajoutés par les systèmes de DRM quand ils s'appliquent aux contenus téléchargés sont les suivants :

- Limitation de la possibilité de lire en le dupliquant le fichier téléchargé sur plusieurs appareils de lecture. Cette limitation repose sur un cryptage du fichier avec des systèmes de clés asymétriques de 128 bits. Les logiciels de lecture sont « activés » sur un appareil avec un compte utilisateur. Les fichiers peuvent être copiés sur d'autres appareils mais ne peuvent être lus que sur un nombre limité de machines, sur lesquelles le même logiciel de lecture est activé avec le même compte.
- Contrôle des droits de copier-coller et d'impression, en termes de nombre de pages, absolues ou par période.
- Contrôle de la durée de lecture autorisée, afin de définir des modèles économiques de prêt.

2.5 Repérages : quelques acteurs de la distribution de livres numériques

L'analyse des modèles économiques présentée s'appuie à la fois sur une analyse théorique des principaux modèles économiques possibles et sur une sélection des modèles économiques effectivement proposés dans le cadre des expériences et des offres commerciales actuelles. La description, sans visée exhaustive, du paysage actuel de quelques acteurs importants de la distribution de livres numériques ne fera qu'illustrer l'analyse générale précédente. Toutes ces pratiques s'inscrivent en effet de manière plus ou moins pure dans l'un des modèles théoriques décrits. Cette description permettra toutefois de prendre conscience du type d'acteurs économiques qui animent le marché des livres numériques et qui pourraient être à des titres divers des partenaires d'*Europeana*, et du dynamisme croissant de ce marché du côté de l'offre. L'angle de présentation choisi est celui de l'offre, plutôt que celui de la demande, les diverses expériences en cours de bibliothèques numériques avec des contenus sous droits pouvant être interprétées comme des applications de ces différents types d'offre, dont la description détaillée serait donc une redondance. Le panorama présenté se concentre exclusivement sur **les offres de distribution de livres numériques** et ne concerne donc ni les sites de revues, ni les portails de recherche, ni les plates-formes d'archivage des fichiers et des métadonnées, comme par exemple la plate-forme « Volltextsuche Online » lancée en 2007 par le Börsenverein, groupement des éditeurs et libraires allemands.

2.5.1 Les librairies de livres numériques

De nombreux sites dans le monde proposent aujourd'hui aux internautes d'acheter des livres numériques récents, selon différents modèles commerciaux, essentiellement ceux décrits plus haut comme L1, L3, L4, L5, L7.

Le modèle commercial le plus répandu est pour le moment le couple (L1, L5), les internautes achetant des licences pérennes titre par titre ou chapitre par chapitre. C'est dans le monde anglo-saxon qu'on compte le plus grand nombre d'offres, mais il en existe aussi dans le reste du monde et notamment en Europe. Ces offres émanent le plus souvent encore de « pure players », qui sont en général des agrégateurs numériques avec des catalogues dépassant fréquemment les 30 000 titres et qui jouent à la fois le rôle de diffuseurs et de distributeurs. On trouve aussi quelques sites d'éditeurs spécialisés ou de librairies en ligne ou traditionnelles qui se fournissent auprès des agrégateurs.

Voici à titre d'exemples non exhaustifs quelques-unes des principales librairies de livres numériques proposant les modèles (L1, L5), illustrant la diversité de cette offre en ligne :

- <http://www.fictionwise.com> (agrégateur, États-Unis)
- <http://www.ebookmall.com> (agrégateur, États-Unis)
- <http://ebooks.palm.com> et <http://www.ereader.com> (agrégateurs, sites d'ebooks spécialisés dans les ebooks pour PDA créés par l'éditeur du logiciel Palm, États-Unis)
- <http://www.franklin.com> (site de Franklin Electronics Publisher, éditeur américain spécialisé dans les dictionnaires électroniques)
- <http://e-literal.com> (agrégateur, États-Unis ; version espagnole : letramovil.com)
- <http://www.ebookimpressions.com> (ebook store pure player , États-Unis)
- <http://www.efollett.com/CategoryDisplay/10001-9603-10051-1?demoKey=d> (site d'ebooks de la librairie américaine efollett, destinée aux étudiants)
- <http://www.powells.com/ebookstore/ebooks.html> (site d'ebooks de la librairie américaine Powells)
- http://www.amazon.com/exec/obidos/tg/browse/-/551440/ref=b_tn_bh_eb/002-1204779-5767200 (site de vente d'ebooks d'Amazon)
- <http://www.ebooks.com> (agrégateur, Australie et États-Unis)
- <http://ebooks.whsmith.co.uk/C76B3580-17A8-4B94-A822-0F7394C11C19/10/46/en/Default.htm> (site d'ebooks de la librairie anglaise Whsmith)
- <http://www.ebookstore.tandf.co.uk/html/index.asp> (site d'ebooks de l'éditeur anglais Taylor and Francis)
- www.beam-ebooks.de (agrégateur, Allemagne)
- <http://www.ciando.com> (agrégateur, Allemagne)
- <http://www.evolutionbook.com> (agrégateur, Italie)
- <http://www.elib.se> (agrégateur émanant d'un regroupement d'éditeurs, Suède)
- <http://ebookstore.ebog.dk> (agrégateur récent ayant négocié un accord avec l'association des éditeurs danois)
- <http://www.eBook.nl> (agrégateur, Pays-Bas)
- <http://www.e-books.com.pl> (agrégateur, Pologne)

- <http://www.digitalia.hu> (agrégateur, Hongrie)
- <http://www.mobipocket.com> (agrégateur français à l'origine, filiale d'Amazon)
- <http://e-books.lavoisier.fr> (site d'ebooks de la librairie Lavoisier, France)
- <http://www.numilog.com> (agrégateur, France)
- ...



Le site d'ebooks de la librairie Whsmith

Certains agrégateurs offreurs proposent également des modèles de librairie numérique de type L3, L4, L7, reposant sur des licences temporaires d'accès aux livres, sous la forme de pay-per-view ou plus souvent d'abonnements. Parmi les librairies déjà citées, c'est notamment le cas du site d'ebooks de l'éditeur Taylor and Francis.

Les sites spécialisés dans ce modèle commercial sont notamment les suivants :

Books 24x7

<http://www.books24x7.com>

Agrégateur américain spécialisé dans la diffusion d'ouvrages techniques et de management. proposés sous forme d'abonnements à des collections. Par exemple : 459\$ par an pour l'abonnement à la collection professionnelle « ITPro ».

Questia

<http://www.questia.com>

Agrégateur américain qui offre plus de 67 000 ouvrages en ligne (et des revues) accessibles par abonnement (par exemple : 20\$ par mois ou 100\$ par an). Ce service s'adresse avant tout aux étudiants qui préparent leurs examens ou rédigent une dissertation.

Cyberlibris

<http://www.cyberlibris.fr>

Agrégateur français qui a créé fin 2006 une offre à destination des particuliers, Cyberlibris Famili, proposant un catalogue d'un millier d'ouvrages sur les thèmes de la vie quotidienne, avec des formules d'abonnement (par exemple : une « formule découverte » avec un engagement de 3 mois, à 7,95 euros TTC par mois).

2.5.2 Les offres de bibliothèques numériques par les agrégateurs

De nombreux agrégateurs proposent des systèmes de bibliothèques numériques, en adoptant la plupart du temps un ou plusieurs des modèles économiques B1, B3, B4 ou B8. Certains de ces agrégateurs se concentrent exclusivement sur le marché des bibliothèques, comme netLibrary, d'autres sont des « agrégateurs purs » qui ne font que de la diffusion auprès des bibliothèques et des librairies, comme Overdrive ou Lightning source, d'autres encore exercent à la fois des fonctions d'agrégation et de distribution directe, comme Fictionwise ou ebooks.com.

Voici à titre d'exemples quelques-unes des principales offres de bibliothèques numériques par des agrégateurs :

Ebrary

<http://www.ebrary.com>

Ebrary fournit des livres, des périodiques et des cartes numériques, provenant des fonds d'une centaine d'éditeurs spécialisés en lettres, sciences sociales, médecine, philosophie, économie, informatique, etc. Les textes sont accessibles uniquement en ligne au format PDF. Selon son site Internet, Ebrary a aujourd'hui plus de 1 000 bibliothèques clientes, partout dans le monde. Le modèle commercial est de type B8 : la bibliothèque paye un abonnement annuel, dont le montant varie en fonction du type et de la taille de l'établissement. Les lecteurs accèdent à l'ensemble des titres proposés par le portail, et le nombre de connexions par ouvrage en ligne est illimité. Ce modèle propose en outre une hybridation avec le modèle de librairie : les lecteurs doivent payer toute impression ou copie du document – à moins que la bibliothèque ne choisisse de régler un abonnement annuel permettant impressions et copies sur une base dépendant du nombre d'utilisateurs (actuellement 0.75\$/an pour chaque étudiant à temps plein).

Libwise

<http://www.libwise.com>

Libwise est l'offre pour les bibliothèques de l'agrégateur Fictionwise. Le modèle économique de Libwise.com est de type B1, un coût d'abonnement à la plate-forme de prêt

s'ajoutant au prix des livres, comme dans quasiment tous les modèles où les achats se font titre par titre.

netLibrary

<http://www.netlibrary.com>

Filiaire d'OCLC, association de bibliothécaires basée au Colorado, netLibrary est le leader mondial de la vente de bibliothèques numériques. Il propose environ 120 000 ouvrages numériques, émanant de tous les principaux éditeurs techniques et académiques anglo-saxons. Son modèle économique principal est de type B1 et B3, avec donc des accès simultanés limités. En plus du prix de chaque ouvrage, netLibrary facture l'accès à chacun des titres, ainsi que les services offerts pour la plate-forme de gestion des accès, correspondant à 15 % du prix du livre pour un accès limité à une période d'un an, et 55 % de ce même prix pour un accès pérenne. Tous les livres numériques sont proposés pour une lecture en ligne, quelques-uns en téléchargement. Plusieurs milliers de bibliothèques sont clientes.

Overdrive – Content Reserve

<http://www.contentreserve.com>

Overdrive est l'un des principaux agrégateurs américains (plus de 40 000 titres), qui alimente en ebooks de très nombreuses librairies, à travers son offre Content Reserve. Il propose également un système de bibliothèque numérique basé sur les modèles B1 et B3, avec des accès limités, titre par titre.

Safari Books Online

<http://safari.oreilly.com>

Né d'une alliance entre les éditeurs O'Reilly Media et Pearson Technology Group, Safari propose un choix de titres techniques et informatiques de plusieurs éditeurs spécialisés auxquels on accède à distance par reconnaissance des adresses IP. La lecture des livres se fait en ligne, avec un logiciel de lecture Safari. Le modèle commercial, assez adapté pour cette catégorie d'ouvrages, est de type B4 : les accès simultanés sont limités, mais partagés entre plusieurs titres d'une collection interchangeables entre eux. Sur le plan marketing, ce modèle est présenté astucieusement comme un système de « jetons » ou de « points », le nombre de points remplaçant les valeurs monétaires des livres. La tarification est effectuée en fonction du nombre de points et du nombre d'accès. La bibliothèque peut modifier à tout moment les titres de son catalogue, ce système de points facilitant l'échange de titres. Elle peut ainsi acquérir les toutes dernières éditions disponibles et surtout changer tous les mois les titres du catalogue si elle le souhaite.

Cyberlibris

<http://www.cyberlibris.fr>

Cet agrégateur français utilise le même modèle commercial (B8 : abonnement à des accès illimités) et la même technologie que Ebrary et adapte le portail « Ebrarian » au site Internet de la bibliothèque. Son catalogue est surtout composé de plusieurs milliers de livres d'économie et de gestion et s'adresse donc en priorité aux Ecoles de commerce. Des services annexes sont proposés, comme des blogs d'auteurs et le partage de commentaires sur les ouvrages.

Numilog

<http://www.numilog.com>

Numilog, auteur de la présente étude, est un agrégateur français qui propose, outre son offre de librairie numérique, et la possibilité d'alimenter d'autres librairies avec son catalogue, une offre de bibliothèques numériques à destination des bibliothèques de lecture publique, universitaires, et des grandes écoles. Cette offre permet aux bibliothèques de prêter des livres numériques à leurs abonnés, en consultation sur place ou en accès à distance sur un site de bibliothèque numérique personnalisé pour chaque établissement. Son catalogue général s'élève à 35 000 titres dans tous les domaines, dont environ 15 000 disponibles pour les bibliothèques. Son modèle commercial est de type B1 et B3, avec des accès limités, titre par titre, auquel s'ajoute un abonnement à la plate-forme technique.

2.5.3 Les offres numériques directes des éditeurs

Un certain nombre d'éditeurs proposent directement des offres de livres numériques aux bibliothèques et établissements d'enseignement. Ces offres ont en général la forme B8 d'abonnement à des collections en accès illimité en ligne, souvent associé à des « big deals » dépendant du nombre d'utilisateurs. Ces éditeurs collaborent également avec des agences d'abonnement comme EBSCO (www.ebsco.com) et SWETS (www.swets.com) pour la distribution de ces offres, et avec des agrégateurs pour la distribution d'offres sous des modèles commerciaux titre par titre, de type L1, B1 et B3. Il s'agit surtout d'éditeurs techniques et universitaires anglo-saxons. Pour la plupart, la production et la distribution de livres numériques fait désormais partie intégrante de leur stratégie concurrentielle, quasiment toutes les nouveautés publiées étant disponibles simultanément sous la forme d'ebook. **Il s'agit d'une « fenêtre d'exploitation » de leur production éditoriale qui n'est plus du tout considérée comme expérimentale, mais comme l'un des axes importants de leur activité courante.** D'où des catalogues disponibles de plusieurs dizaines de milliers d'ebooks. Les tarifs des abonnements ne sont pas souvent publics car ils font l'objet de négociations qui peuvent dépendre du pays et de l'historique d'achat de l'établissement avec l'éditeur concerné, avec des clauses de confidentialité.

On peut citer à titre d'exemples, dominant le marché des offres directes, les quatre plus grands groupes d'édition scientifique et technique mondiaux :

Elsevier

<http://www.elsevier.com>

Le groupe Elsevier est le premier groupe mondial d'édition scientifique, en particulier pour les revues. Il propose en direct des abonnements à des collections thématiques d'ouvrages numériques, en particulier dans les domaines de l'environnement et de la médecine.

Springer

<http://www.springer.com>

Springer propose des abonnements à des collections de plusieurs milliers de livres numériques en sciences. Le système d'abonnement fonctionne de la même manière que pour ses revues en ligne. Par le biais du consortium Couperin, Springer propose notamment aux bibliothèques françaises un prix d'abonnement – non public – négociable selon le nombre

de participants à la Springer eBook Collection qui regroupe 12 bibliothèques thématiques d'ouvrages numériques dans le domaine Science-Technologie-Médecine.

Wiley

<http://www3.interscience.wiley.com>

Wiley propose des collections d'ebooks dans trois disciplines : chimie, médecine et sciences de la vie, technologies de l'information et de la communication. Il propose également un ensemble d'ouvrages de référence (reference works), tels que des guides, des encyclopédies et la base de données Express Exec (<http://www.expressexec.com>) qui diffuse une centaine de livres traitant des concepts fondamentaux de la gestion. Wiley InterScience® propose en outre une collection de e-books personnalisés, créée à partir de plus de 1 600 ouvrages numériques disponibles dans la Wiley OnlineBooks™. Le modèle commercial est toujours celui d'abonnements annuels permettant des accès illimités à des livres lus en ligne.

Taylor and Francis

<http://www.ebookstore.tandf.co.uk>

Taylor and Francis dispose du plus ancien et plus important catalogue de livres numériques pour un groupe d'édition, dépassant les 60 000 références et vendu sous des formes très variées. Il propose notamment aux bibliothèques, sous forme d'abonnements, la Taylor and Francis eCollections, composée de 10 bouquets contenant plus de 4 000 titres dans les domaines scientifique et des sciences humaines et sociales.

3 Recommandation : le modèle économique proposé

Une question fondamentale pour *Europeana* est de savoir quel(s) type(s) de fonction elle devrait assurer en propre parmi les positions fonctionnelles identifiées, pour les contenus sous droits, et lesquelles doivent être traitées en collaboration ou routées vers d'autres acteurs spécialisés. Autrement dit, sur quel(s) carré(s) du schéma organisationnel du §2.2 le projet se positionne-t-il ? *Europeana* doit-elle être :

- une plate-forme d'hébergement des fichiers numériques et des métadonnées ?
- un portail de recherche ?
- une bibliothèque numérique ?
- une librairie numérique ?
- un diffuseur-agrégateur ayant des liens commerciaux avec des bibliothèques numériques et des librairies numériques ?

L'essentiel de la réponse à la question du bon modèle économique est une réponse à cette question fonctionnelle.

Afin de répondre de manière pertinente à cette question, nous nous appuierons :

- sur les analyses des modèles économiques possibles et pratiqués pour vendre des livres numériques déjà présentés dans le §2 ;
- sur l'observation des fonctionnalités du site *Europeana* <http://www.europeana.eu> ;

- sur la liste des contraintes formées par les objectifs respectifs des deux principales parties prenantes : la Bibliothèque Nationale de France d'une part, les éditeurs d'autre part.

3.1 Les fonctionnalités du site Europeana

Un prototype du site *Europeana* a été ouvert le 23 mars 2007 à l'adresse www.europeana.eu.

Le site rassemble à cette date 12 000 documents du domaine public en provenance de fonds français, hongrois et portugais. L'objectif affiché par la BnF est d'ajouter dans les mois qui viennent des documents issus de l'« océrisation » d'une grande partie de la base de 80 000 documents de *Gallica* et environ 100 000 documents numériques nouveaux par an.

Les fonctionnalités de recherche proposées sur le site combinent :

- la recherche plein texte dans la totalité du corpus ;
- la recherche par thème ou catégorie de documents, en particulier à travers des filtres de langue, époque de publication et provenance.

La recherche plein texte est effectuée pour le moment grâce au moteur de recherche open source LUCENE, et devrait dans une phase ultérieure être complétée par un moteur de recherche sémantique. Elle permet de présenter les documents classés par ordre de pertinence, le degré de pertinence étant indiqué à côté du document trouvé. L'avantage de cette solution de recherche est d'une part que les critères de pertinence sont publics (le moteur de recherche étant libre et analysable par tous) et d'autre part qu'ils sont combinés à des filtres de recherche permettant une démarche de recherche multilingue et intelligente dans les corpus.

Pour chaque document trouvé, il est notamment possible de :

- visualiser un bref résumé, les termes de la recherche trouvés étant surlignés ;
- visualiser les métadonnées associées à chaque ouvrage ;
- ajouter la notice du document à « Ma bibliothèque / Mes documents » où chaque internaute, après une authentification, pourra retrouver rapidement ses œuvres favorites ;
- avoir accès, pour les contenus hors droits présentés (dans le domaine français pour le moment), au texte intégral du livre avec un surlignage dans le texte des occurrences de l'expression recherchée.

Les textes sont affichés en mode image avec texte sous image, avec de nombreuses fonctionnalités de lecture. Il est notamment possible :

- de sauvegarder le texte ;
- de le télécharger, en mode image ou en mode texte, mais sans DRM ;
- de l'imprimer ;
- de l'envoyer par e-mail ;
- d'ajouter des « étiquettes » pour repérer des pages.

Dans une prochaine version du site, il est prévu d'ajouter des fonctionnalités de commentaires et de partage de ces commentaires avec d'autres utilisateurs.

L'ensemble de ces fonctionnalités est tout à fait ergonomique et « orienté utilisateur ». Bien entendu, toutes ne sont pas extensibles en l'état aux contenus sous droits pour les raisons de protection déjà évoquées.

Voici un exemple de parcours de navigation sur le site *Europeana* actuel :



La page d'accueil d'Europeana

The screenshot shows a web browser window displaying search results on the Europeana website. The search term 'livre' is entered in the search bar. The results page shows 10/6901 results. Three results are visible:

- Result 1:** Aide-mémoire du libraire et de l'amateur de livres : répertoire d'ouvrages rares ou curieux en tous genres, anciens et modernes, éditions originales, livres à gravures des XVIe, XVIIe, XVIIIe et XIXe siècles, impressions rares, etc., avec l'indication de leur valeur dans le commerce. 2. Lal-Z. 1907 / par un ancien libraire - 1906. Extrait: Éditeur : aide-mémoire du libraire et de l'amateur de livres (Paris) Sujet : Livres rares - Bibliographie 100 %
- Result 2:** Aide-mémoire du libraire et de l'amateur de livres : répertoire d'ouvrages rares ou curieux en tous genres, anciens et modernes, éditions originales, livres à gravures des XVIe, XVIIe, XVIIIe et XIXe siècles, impressions rares, etc., avec l'indication de leur valeur dans le commerce. 1. Première partie. A-Lal. - 1906 / par un ancien libraire - 1906. Extrait: Éditeur : aide-mémoire du libraire et de l'amateur de livres (Paris) Sujet : Livres rares - Bibliographie 100 %
- Result 3:** Les travailleurs du livre et du journal / par G. Renard - 1925-1926. Titre : Les travailleurs du livre et du journal / par G. Renard. Auteur : Renard, Georges (1847-1930). Éditeur : G. Doin (Paris). Date : 1925-1926. Sujet : Livres - Industrie et commerce - France. Type : texte, monographie en plusieurs volumes. Langue : Français. Format : application/pdf. Droits : domaine public. Identifiant : ark:/12148/bpt6k20843r. Provenance : Bibliothèque Nationale de France. Extrait: Sujet : Livres - Industrie et commerce - France.

Résultat d'une recherche sur le terme « livre » dans les contenus en français

The screenshot shows a web browser window displaying the Europeana digital library interface. The browser's address bar shows the URL: <http://www.europeana.eu/ark:/12148/bpt6k208435r.r=livre>. The page title is "Europeana Consultation : Les travailleurs du livre et du journal / par G. Renard - page - NP - - Internet Explorer fourni par".

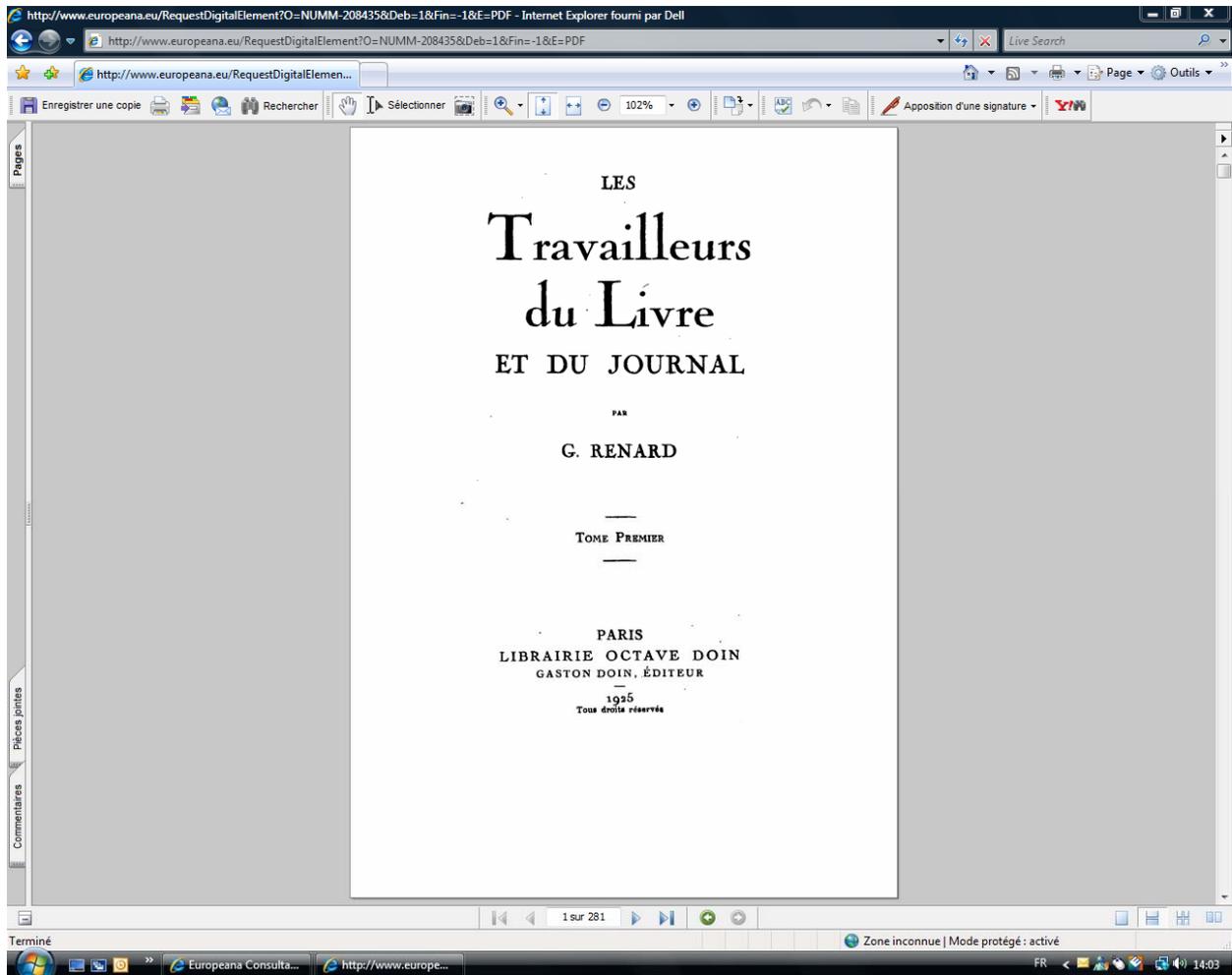
The main content area features the Europeana logo and a search bar containing the word "livre". Below the search bar, there is a navigation bar with the text "Les travailleurs du livre et du journal / par G. Renard / Renard, Georges (1847-1930), 1925-1926" and a language selector set to French. A "Rechercher dans ce document" sidebar on the left shows "161 pages trouvées" and a list of page thumbnails, with "NP" and "p.I" highlighted. A "Ma bibliothèque" sidebar on the right contains a "Se connecter" button and a "Créer un compte" link.

The central viewer displays the title page of the book "LES Travailliers du Livre ET DU JOURNAL" by G. RENARD, TOME PREMIER, published in Paris by Librairie Octave Doin in 1925. The word "Livre" in the title is highlighted in yellow. The viewer includes navigation controls and a zoom level of 100%.

Accès à l'un des ouvrages trouvés avec le « Reader » d'Europeana

The screenshot shows a web browser window displaying the 'Europeana Ma bibliothèque' interface. The browser's address bar shows the URL: <http://agora.europeana.eu/Panier?Opt=espacePersonnel&AppOrigine=Gallica>. The page features the Europeana logo and a search bar. A sidebar on the left contains the user's alias 'numilog' and a 'Modifier' button. The main content area is titled 'Ma bibliothèque' and includes a 'Travailler' button. Under the 'Mes documents' section, there is a list of 'Documents non classés (1)' with the following details for the selected document: **Les travailleurs du livre et du journal / par G. Renard - 1925-1926**. The document's metadata includes: Titre: Les travailleurs du livre et du journal / par G. Renard; Auteur: Renard, Georges (1847-1930); Editeur: G. Doin (Paris); Date: 1925-1926; Sujet: Livres - Industrie et commerce - France; Type: texte, monographie en plusieurs volumes; Langage: Français; Format: application/pdf; Droits: domaine public; Identifiant: ark:/12148/tp6t208435r. Below this, the 'Mes étiquettes' section indicates that no tags are currently present and provides instructions on how to add tags. A right-hand sidebar titled 'Ma bibliothèque' shows a 'Mon bureau' section with links to 'Ma page d'accueil', 'Mes documents', and 'Mes étiquettes', along with a 'Se déconnecter' button. The browser's taskbar at the bottom shows the system tray with the date and time set to 14:08.

Ajout de la notice de l'ouvrage sélectionné à « Mes documents »



Téléchargement de l'ouvrage dans Adobe Reader (sans DRM)

3.2 Les principaux objectifs et réquisits de la BnF pour *Europeana*

Sur la base du cahier des charges de la présente étude, des documents résumant les précédentes études menées pour initier et développer le projet (notamment le « Livre Blanc du Comité de Pilotage » remis en janvier 2006) et des entretiens directs menés avec les responsables du projet, il ressort que les principaux objectifs de la BnF en ce qui concerne l'ajout de contenus sous droits dans *Europeana* peuvent être résumés de la manière suivante :

1. Favoriser l'apport régulier à *Europeana* et en quantité importante de livres sous droits, en priorité dans un certain nombre de secteurs documentaires correspondant à la politique documentaire générale poursuivie par *Europeana* mais également de livres récents dans tous les domaines (voir plus bas le paragraphe sur la politique documentaire).
2. Assurer dans toute la mesure du possible une continuité des pratiques d'accès aux contenus patrimoniaux et sous droits et notamment :
 - Permettre une recherche structurée commune dans les corpus patrimoniaux et sous droits.
 - Permettre dans les deux cas un accès aux textes intégraux.
 - Permettre à l'internaute de bénéficier de certains outils de navigation communs (comme l'ajout à « Mes documents »).
3. Respecter le droit de la propriété intellectuelle sur tous les ouvrages présentés.
4. Permettre autant que faire se peut à tous les acteurs professionnels de la chaîne du livre traditionnelle ou de la nouvelle chaîne du livre numérique de jouer leur rôle.
5. Ne pas jouer de rôle commercial direct.
6. Ne pas demander d'exclusivité sur les contenus présentés dans *Europeana*.
7. Eviter de devoir payer des coûts d'acquisition directs importants aux éditeurs.
8. Ne demander aucune exclusivité pour l'apport des ouvrages numériques sous droits à *Europeana*.

Ces objectifs ne sont pas a priori contradictoires, mais ne sont pas directement faciles à concilier, en particulier la volonté de susciter une collaboration massive des éditeurs, sans prévoir de budget de *Europeana* pour les y inciter ni de rôle commercial direct. Certains de ces objectifs généraux méritent par ailleurs d'être spécifiés de manière plus précise. L'apport de contenus sous droits doit être précisé dans sa nature : s'agit-il d'ajouter des

métadonnées, des possibilités de recherche plein texte ou des possibilités de lecture du texte intégral et dans quelles conditions ? Le respect des droits des ayants droit et du rôle des professionnels de la chaîne du livre doit être précisé dans son mode de fonctionnement et dans son extension.

3.3 Les principaux objectifs et réquisits des éditeurs pour participer à Europeana

Sur la base des entretiens menés avec plusieurs éditeurs, en particulier en France, et avec des organisations qui les représentent, en France et en Europe, les principaux objectifs et réquisits des éditeurs pour collaborer à *Europeana* peuvent être résumés de la manière suivante :

1. Promouvoir leurs fonds éditoriaux au sein d'un moteur de recherche international à forte visibilité et proposant une hiérarchisation intelligente et « non commerciale » de ces fonds, liés exclusivement à leur pertinence pour la recherche des internautes.
2. Rentabiliser l'accès aux textes intégraux en ligne ou par téléchargement, en étant normalement rémunérés pour ces accès et en évitant le risque de déséquilibrer le modèle économique actuel de la filière livre.
3. Conserver la maîtrise de la production, de l'archivage et des usages de leurs fonds numériques, en limitant au maximum les transferts de fichiers non protégés vers des plates-formes tierces.
4. Garantir le respect des droits d'auteur sur les livres diffusés.
5. Permettre autant que faire se peut que les libraires, partenaires indispensables des éditeurs, bénéficient également des effets potentiels de promotion des ouvrages permis par *Europeana* sans que la diffusion numérique ne devienne une menace pour leur activité.
6. Bénéficier si possible d'aides publiques à la numérisation rétrospective de leurs fonds, en particulier pour leurs fonds plus anciens, c'est-à-dire lorsque les investissements dans cette numérisation présentent une visibilité moindre de retour financier à court terme.

On peut remarquer que la BnF et les éditeurs partagent les objectifs de respect des ayants droit et des autres professionnels de la chaîne du livre. Les éditeurs ne peuvent qu'agrée également à l'objectif de non exclusivité. La conciliation des autres objectifs méritent une analyse comparative des modèles économiques envisageables.

3.4 Analyse des modèles économiques envisageables et recommandation

Afin de déterminer le modèle économique recommandé pour *Europeana*, nous proposons de raisonner en présentant les avantages et inconvénients des différents modèles économiques envisageables et en éliminant ceux qui ne paraissent pas souhaitables. L'application de cette méthode conduit à sélectionner résolument un des modèles économiques possibles, mais en envisageant des variantes possibles. Ce modèle économique conduira à une réponse précise à la question de la position fonctionnelle que devrait avoir *Europeana*, pour les contenus sous droits, dans la chaîne du livre.

Le **premier modèle économique** envisageable est celui d'un accès gratuit aux textes numériques financé exclusivement par de la publicité et par la promotion des ouvrages papier. L'avantage de ce modèle pour les utilisateurs n'a pas besoin d'être développé. Mais ce modèle, pour des raisons déjà évoquées, n'est à l'heure actuelle ni réaliste ni acceptable. Les financements publicitaires trouvés ne seraient certainement pas suffisants pour rémunérer les ayants droit, en particulier si on considère les axes prioritaires de la politique documentaire, qui privilégie les ouvrages de référence à contenu de connaissance et culturels sur les seuls « best-sellers » du moment ou les guides de voyage. Ce modèle est très inhabituel pour le marché des livres et se heurterait sans aucun doute, mis en avant sur une telle échelle, au rejet aussi bien de nombreux éditeurs que des auteurs, libraires, bibliothécaires et même de beaucoup d'usagers. Le classement des livres, comme il est reproché à d'autres portails de recherche privés, serait trop susceptible d'être affecté par ces (très éventuelles) ressources publicitaires. Pour toutes ces raisons, il n'est donc pas souhaitable à notre avis que ce modèle soit utilisé, même comme modèle complémentaire, par un portail public tel qu'*Europeana*.

Si on élimine le modèle publicitaire, qu'il fallait tout de même considérer pour des raisons d'exhaustivité de l'analyse, les deux catégories de modèles économiques possibles sont celles présentées au §2.3 :

- le modèle de bibliothèque numérique
- le modèle de librairie numérique

Le **modèle de bibliothèque numérique** présente l'avantage d'être a priori le plus « naturel »... pour un projet émanant de bibliothèques. *Europeana*, et les bibliothèques nationales qui constitueront son fonds d'ouvrages, pourraient effectuer la médiation économique propre aux bibliothèques afin de financer la création progressive d'un fonds d'ouvrages numériques sous droits proposés en accès gratuit aux citoyens d'Europe. Ce qui représenterait une très belle ambition pour un premier genre de « service public culturel européen ». Ce qui représenterait un avantage indéniable pour les internautes lecteurs, comparable à celui du modèle publicitaire. Toutefois, l'usage de ce modèle à l'échelle du projet *Europeana* poserait de très sérieux et quasi incontournables problèmes. Pour s'en rendre compte, il suffit de considérer les principales sous-catégories de modèles économiques de bibliothèques numériques logiquement et pratiquement possibles :

- Si on considère d'une part les **modèles à accès simultanés limités**, la question qui se pose est la suivante : combien d'accès simultanés est-il pertinent de proposer pour per-

mettre un usage « normal » des ressources d'*Europeana* par tous les internautes européens ? Dans la mesure où il s'avèrerait difficile, tant techniquement que politiquement, de fermer *Europeana* aux internautes du reste du monde, il faut élargir cette question en se demandant combien d'accès simultanés sont pertinents pour un usage par tous les internautes du monde ? Il semble quasiment impossible de répondre à cette question. Ou alors il faudrait proposer des chiffres bien trop élevés, et de toute façon très arbitraires, pour qu'un financement public en soit acceptable dans le contexte actuel des fonds disponibles tant au niveau national qu'europpéen.

A titre de simple illustration chiffrée de cette difficulté, considérons le scénario suivant :

- *Europeana* décide de proposer 100 accès simultanés à tous les livres sous droits – ce qui est très peu si on considère la population mondiale d'utilisateurs potentiels – et les risques économiques objectifs pris par les éditeurs si une mutualisation des accès numériques d'une telle ampleur avait lieu, même en contrôlant les accès simultanés. En effet, la granularité actuelle de l'implantation géographique des bibliothèques est en soi un facteur d'achat de plus d'exemplaires dont, sinon, les accès pourraient être mutualisés. En particulier pour les livres les plus spécialisés.
- chaque année, environ 35 000 nouveaux livres sont proposés par les éditeurs en France (sans compter les réimpressions). Supposons que soit apportée à *Europeana* seulement la moitié de ces « pures nouveautés », soit 17 500 titres par an.
- le prix moyen du livre est de 10€H.T.
- la TVA sur les livres numériques est 19,6%.
- le budget d'*Europeana* pour proposer ces accès simultanés serait alors de près de 21 Millions d'euros par an... pour les seuls livres français.

Ce calcul n'inclut pas le budget nécessaire pour financer les accès simultanés aux fonds d'ouvrages antérieurs aux nouveautés, qu'*Europeana* souhaite ajouter à son catalogue, et qui comprend cette fois plusieurs centaines de milliers de titres. Financer l'accès sur les mêmes bases à la moitié de 20 années de production éditoriale moyenne représenterait un budget de plus de 400 Millions d'euros !

Compte tenu du point 7 de la liste des objectifs de la BnF, un tel modèle n'est donc absolument pas envisageable pour le moment... même en divisant par un facteur 10 les accès simultanés par rapport au scénario présenté. Il est très peu probable, vu les montants évoqués, qu'il le soit, en tout cas à cette échelle, par aucune des bibliothèques nationales européennes partenaires potentielles d'*Europeana*.

- Si on considère d'autre part les **modèles à accès illimités**, la question qui se pose est la suivante : quel est la valeur du « big deal » acceptable par les éditeurs français, puis par leurs confrères européens, pour permettre un accès illimité des internautes du monde entier à la totalité des ouvrages qu'ils apporteront à *Europeana* ? Quelle négociation collective permettrait de les dédommager à une hauteur raisonnable de cet usage ? Si on considère à nouveau que le projet d'*Europeana* est de référencer une partie importante des futures nouveautés éditoriales ainsi que, progressivement, de très nombreux ouvrages sous droits des 70 dernières années, il est impossible de répondre à cette question. Ou plutôt, la seule réponse peut-être acceptable par les éditeurs – et les libraires ? – serait un budget annuel

proche de 100% de la totalité de la valeur annuelle du marché du livre ! Il n'est pas besoin d'insister sur l'absurdité d'un tel calcul et d'une telle éventualité.

- Si on considère enfin le **modèle économique « hybride »** dans lequel la bibliothèque paye pour chaque accès d'un lecteur à un livre, on aboutit à une autre impasse : soit *Europeana* définit ex ante un budget plafond, qui risque, vu l'ampleur des usagers potentiels, d'être vite atteint et de conduire à un arrêt des services d'accès offerts jusqu'à l'année suivante, soit ce modèle aboutit à un mécanisme dans lequel le budget annuel ne peut pas être défini à l'avance et risque, ex post, d'aboutir à un montant explosif inacceptable.

Par conséquent, aucun des modèles économiques envisageables de type « bibliothèque numérique » ne fournit de réponse raisonnable et potentiellement acceptable pour la rémunération des ayants droit pour l'accès à leurs œuvres dans le cadre d'*Europeana*. Ceci vient bien sûr d'une part du fait du point 7 des objectifs de la BnF, qui ne dispose pas de budgets significatifs pour ce faire. Mais plus profondément, et même si d'autres bibliothèques nationales ou gouvernements disposaient de dotations budgétaires plus importantes, cela résulte de la taille de la collectivité d'usagers potentiels : les internautes d'Europe, voire du monde entier. Aucune bibliothèque numérique ne s'adresse à une telle communauté et ce changement d'échelle radical conduit à un changement de nature : un projet de bibliothèque numérique « universelle » (en termes de contenus) ne nous semble tout simplement pas envisageable aujourd'hui à cette échelle.

Il ne reste par conséquent qu'à envisager pour *Europeana* un **modèle économique de librairie numérique**, dans lequel ce sont les usagers qui paient directement pour l'accès aux livres de leur choix. Ce type de modèle peut se heurter aux deux objections sérieuses suivantes, du côté des bibliothèques :

- a. L'idée qu'*Europeana* propose un accès payant aux usagers pourrait sembler contraire à la philosophie habituelle d'une bibliothèque et à l'idée d'une « bibliothèque » numérique européenne. Plusieurs bibliothèques nationales, en particulier dans les pays scandinaves, risquent de considérer cette idée opposée à leur conception de leur rôle traditionnel, voire même à leurs propres statuts.
- b. L'objectif 5 de la BnF est de ne jouer aucun rôle commercial. Comment cet objectif est-il conciliable avec un modèle économique de librairie numérique ?

Une réponse à la première objection est possible si on répond de manière adaptée à la question fonctionnelle : quelle position fonctionnelle *Europeana* devrait-elle adopter dans la chaîne du livre, et en particulier du livre numérique ? On peut fournir la réponse suivante :

- *Europeana* devrait être pleinement une « bibliothèque numérique » pour les contenus patrimoniaux hors droits, c'est là son rôle en tant qu'institution regroupant potentiellement les bibliothèques nationales des différents pays européens.

- Mais *Europeana* pourrait être essentiellement un **« portail de recherche » pour les contenus sous droits**. Cela n'est pas opposé au rôle normal des bibliothèques, qui, de plus en plus à l'ère numérique et des « bibliothèques hybrides », jouent également un rôle d'organisation intelligente de l'accès à l'information et de signalement des ressources numériques disponibles, qu'elles ne peuvent pas toutes acheter sur leurs fonds propres. Cela représente un nouveau type de service au public. Et si *Europeana* joue le rôle de portail de recherche intelligent pour les contenus sous droits, et l'assume en tant que tel dans la

définition de ses missions, il n'y a plus de contradiction, déontologique ou institutionnelle, avec l'idée qu'une bibliothèque ne fait pas payer les lecteurs pour l'accès à ses fonds : il ne s'agit pas là de « ses » fonds.

Une réponse à la deuxième objection est également possible si le rôle commercial impliqué par le modèle de librairie numérique n'est pas joué par *Europeana*, mais par des partenaires privés associés à *Europeana*. En tant que portail de recherche, *Europeana* peut simplement renvoyer vers ces partenaires pour gérer la commercialisation de l'accès payant aux contenus sous droits. Dans ce cas, toutes les conditions sont remplies pour satisfaire au mieux la totalité des objectifs présentés de la BnF et des éditeurs, notamment :

- l'incitation des éditeurs à fournir, sur des bases non exclusives, des quantités importantes de livres à *Europeana* afin qu'ils y soient signalés et qu'ils puissent les vendre selon des modèles économiques et des solutions de protection des droits d'auteurs qui leur conviennent, avec des systèmes DRM adaptés ;
- la possibilité pour *Europeana* de proposer au sein de son moteur de recherche une information très complète et à visée progressivement exhaustive sur les ouvrages européens, dans un contexte de respect des droits d'auteur, tout en offrant aux internautes une solution pour accéder aux textes intégraux, sans avoir à la financer sur fonds publics et sans jouer de rôle commercial.

En conclusion, nous recommandons à la BnF d'utiliser pour l'accès aux contenus sous droits dans *Europeana* le modèle économique de librairie numérique.

Ce modèle présente de très nombreux avantages et nous apparaît le mieux adapté pour répondre aux exigences posées par la BnF dans le cadre du cahier des charges de la présente étude.

Notre recommandation implique qu'*Europeana* se positionne de manière « hybride » à la fois comme une bibliothèque numérique au sens fort pour les contenus du domaine public et comme un portail de recherche pour les contenus sous droits.

Nous recommandons par ailleurs que l'utilisation des fonctionnalités de recherche demeure toujours gratuite, quel que soit le type de contenus recherché.

Cette offre de recherche gratuite correspond à la fois à la philosophie adoptée par les principaux sites privés proposant des moteurs de recherche plein texte tels que *Google Book Search* ou *Amazon Search Inside* et à la démarche attendue d'une bibliothèque publique.

Malgré la gratuité du service offert aux usagers, la fourniture par les éditeurs à *Europeana* des éléments techniques (fichiers, métadonnées) permettant à la recherche de s'effectuer en mode plein texte et de manière « intelligente », a **une indéniable valeur sociale et économique** pour le service qu'elle rend au public et devrait être reconnue comme telle dans le dialogue partenarial entre *Europeana* et les éditeurs. L'autorisation d'intégration dans un moteur de recherche plein texte est par ailleurs **loin d'être neutre quant au droit moral sur les œuvres**, point auquel les auteurs notamment sont très sensibles. Les contreparties de cet apport, si elles ne sont pas censées être de l'ordre d'une contrepartie financière

directe, devraient à tout le moins résider dans les espoirs de promotion des ouvrages sous droits, auxquels *Europeana* devra montrer qu'elle porte une attention, et, pour certains d'entre eux, dans de possibles aides incitatives à leur numérisation.

Europeana, qui pourra héberger certains fichiers pour le fonctionnement de son moteur de recherche (voir plus bas), n'a pas vocation à être une plate-forme d'hébergement des fichiers numériques des éditeurs, qui souhaitent conserver la pleine maîtrise de cet hébergement, les fichiers des textes intégraux n'étant accessibles au public qu'à partir d'autres plates-formes. Elle n'a pas non plus vocation à jouer un rôle commercial d'agrégateur pour d'autres opérateurs de distribution des livres numériques.

Le fonctionnement précis de ce modèle économique doit encore être précisé, en particulier à travers les points suivants :

- la nature des partenaires privés envisagés pour gérer le modèle de librairie numérique et le choix de ces partenaires, question qui pose celle de l'intégration des libraires ou autres partenaires dans le schéma de diffusion considérée ;
- les modalités d'association des auteurs et de leurs ayants droit à ce schéma ;
- le parcours de navigation concret des internautes impliqué par ce schéma ;
- les modalités techniques de fonctionnement et leurs implications sur la numérisation des livres.

Par ailleurs, certaines variantes autour de ce modèle ou perspectives de développement pourront aussi être considérées afin qu'*Europeana* soit un projet dynamique, souple et ouvert d'une part aux cultures et conceptions propres à d'autres pays européens, d'autre part aux évolutions technologiques permanentes caractéristiques de l'économie numérique.

3.5 Le modèle économique recommandé : quels partenaires ?

3.5.1 Les opérateurs de diffusion agréés

Il est concevable d'organiser l'accès contrôlé (pour le parcours gratuit de pages trouvées à l'issue d'une recherche plein texte) et payant (pour l'accès aux textes intégraux) aux contenus numériques sous droits selon plusieurs modalités :

1. **la concession de l'exploitation d'une librairie numérique unique** labellisée *Europeana* par la BnF dans un premier temps, par un groupe de bibliothèques nationales ou par les institutions européennes dans un deuxième temps, à un opérateur unique, choisi par une procédure d'appel d'offre.
2. **la conclusion de partenariats directs** avec plusieurs opérateurs de diffusion privés, apportant au moteur de recherche d'*Europeana* des données sur les fonds numériques dont les éditeurs leur ont confié l'exploitation, en contrepartie d'une redirection des usagers vers leurs plates-formes pour l'accès contrôlé à ces fonds ;

3. **des liens vers les différents sites**, d'éditeurs, d'agrégateurs, de librairies en ligne, où les livres numériques sont commercialisés.

Nous recommandons, au démarrage d'Europeana, de privilégier la solution n°2 de conclusion de partenariats directs avec plusieurs opérateurs de diffusion des accès aux contenus sous droits, spécialisés dans la diffusion numérique sécurisée.

En effet :

- Compte tenu de l'organisation de l'édition européenne, très largement sur des bases nationales, et de la volonté des éditeurs de préserver leur indépendance et de maîtriser l'archivage et les modalités de commercialisation de leurs fonds numériques, seuls ou avec les partenaires de leur choix, la solution n°1 présenterait un caractère trop monolithique et risquerait de présenter un très faible degré d'acceptation. En France par exemple, il est peu probable qu'un groupe éditorial tel que le groupe Hachette accepte d'être diffusé par un opérateur privé tiers imposé par *Europeana*, qui aurait dès lors un accès à ses fonds numériques et à la gestion de ses clients. Aucun opérateur privé n'aurait probablement la capacité de convaincre la totalité des éditeurs européens de lui confier la gestion commerciale de leurs ouvrages numériques, étant donné la position dominante que cela lui conférerait.

- La solution n°3 présente, au moins en phase de démarrage, un risque de complexité trop importante quant au parcours des internautes depuis *Europeana*, nuisant à l'ergonomie de l'ensemble. Par ailleurs, elle laisse ouverte la question du choix de l'opérateur qui fournirait à *Europeana* les éléments techniques (fichiers pour la recherche plein texte, métadonnées et notices des livres), nécessaires pour l'intégration des contenus dans son moteur de recherche sur des bases homogènes et celle du respect de normes déontologiques et de protocoles d'échanges de données avec *Europeana* (pour l'authentification des usagers par exemple).

- La solution n°2 présente l'avantage de concilier souplesse et possibilité d'harmonisation : chaque éditeur pourrait choisir l'opérateur qui a sa confiance pour apporter ses ouvrages numériques au portail de recherche *Europeana* et gérer les accès contrôlés aux textes des ouvrages trouvés. Pour autant, les internautes seraient dirigés en première instance vers une plate-forme unique pour chaque livre afin d'en parcourir selon diverses modalités les contenus intégraux. Ces opérateurs privés, choisis par les éditeurs, devraient faire l'objet d'un agrément préalable d'*Europeana* afin de satisfaire un cahier des charges technique et déontologique (voir plus bas).

Nous recommandons d'utiliser une procédure de double choix des opérateurs de gestion des accès aux livres sous droits : par les éditeurs d'une part, sur la base d'un opérateur unique associé à chaque ouvrage, par *Europeana* d'autre part, sur la base d'une procédure d'agrément contractuelle.

Les opérateurs privés choisis selon cette procédure seront responsables de :

- **la gestion des accès contrôlés aux pages trouvées** à l'issue d'une recherche sur *Europeana*, avec des logiciels de lecture en ligne intégrant des mesures de protection appropriées pour les éditeurs ;

- **la gestion des accès payants aux textes intégraux** des livres numérisés selon divers modèles économiques de librairie numérique. Il est recommandé que les opérateurs proposent dans toute la mesure du possible **une gamme variée de solutions de paiement** afin de répondre à la diversité des besoins des internautes et aux catégories différentes d'utilisateurs (grand public, chercheurs et enseignants, laboratoires et institutions...). En particulier, **il serait souhaitable de proposer un choix entre les trois modèles économiques L1, L3 et L4** : achat de livres à titre pérenne, consultation pendant une durée limitée d'un ouvrage, abonnement à une collection de livres. Dans le cas de l'achat, l'internaute pourrait avoir le choix entre un accès pérenne aux livres numériques en ligne ou un téléchargement sur un appareil de lecture, en fonction des formats de fichiers proposés par l'opérateur pour ces livres. Dans tous les cas, l'opérateur devrait faire son affaire de la prise de commande, de l'encaissement, de la livraison, de la sécurisation des accès ou des fichiers téléchargés avec les technologies de DRM appropriées, du paramétrage de ces technologies conformément à leurs contrats avec les éditeurs et du service après-vente pour les clients.

Il ne devrait pas être de la responsabilité d'*Europeana* de définir l'équilibre économique entre ces trois modèles et le niveau de tarif retenu. Celui doit être déterminé seulement par les opérateurs commerciaux, en relation avec les éditeurs qu'ils représentent. Toutefois, ***Europeana* pourrait demander aux opérateurs agréés de respecter certains principes communs**, par exemple :

- la proposition des 3 modèles de tarification L1, L3, L4 ;
 - un prix homogène pour la consultation pendant une durée donnée ;
 - des prix de souscription qui évitent autant le « dumping » (à l'envers des « big deals », afin d'attirer les clients et de créer des habitudes et des effets de réseaux et présente l'inconvénient de ne pas être un modèle durable dans le temps) que les tarifs prohibitifs.
- Cette « harmonisation minimale » peut être justifiée au nom de considérations pratiques et déontologiques. Elle n'est toutefois envisagée qu'à titre de suggestion, et méritera une discussion avec les éditeurs, lesquels naturellement sont susceptibles de préférer une totale liberté de tarification.

Compte tenu des positions fonctionnelles au sein de la « chaîne du livre numérique » présentées au §2.2, **il semble naturel que les opérateurs agréés pour cette gestion des accès aux livres sous droits soient des diffuseurs numériques, ou « agrégateurs »**. Ces acteurs, spécialisés dans la distribution et la diffusion sécurisée de contenus numériques, disposant déjà d'accords contractuels avec les éditeurs pour la diffusion et de catalogues de fichiers numériques, sont les plus à même, et dans des délais rapides, de satisfaire au cahier des charges technique et commercial des opérateurs agréés, et d'apporter à *Europeana* des données sur les fichiers numériques qu'ils exploitent déjà. Il existe aujourd'hui en Europe suffisamment de diffuseurs numériques opérationnels pour jouer ce rôle et ainsi permettre au modèle économique recommandé de se mettre en place dès les prochains mois. Leur nombre devrait encore s'accroître dans les années à venir, qu'il s'agisse de diffuseurs indépendants ou créés par des groupes d'éditeurs, sans qu'il ait vocation, par nature, à devenir

aussi important que celui des points de distribution finaux. Ces diffuseurs peuvent donc jouer auprès d'*Europeana* la même fonction économique et mathématique qu'ils jouent auprès des éditeurs et des distributeurs : la réduction du nombre d'accords bilatéraux nécessaires.

Toutefois :

- Chaque éditeur ou groupe d'éditeurs devrait pouvoir, s'il le souhaite, demander un agrément direct à *Europeana* pour être son propre diffuseur numérique et gérer lui-même l'exploitation commerciale de ses livres numériques. Il est extrêmement probable que seuls quelques grands groupes souhaiteront jouer ce rôle en direct.

- Les opérateurs agréés, gérant l'accueil des usagers en provenance d'*Europeana* et l'accès contrôlé aux pages trouvées, devraient prévoir, autant que faire se peut, **une redirection des usagers, devenus clients, vers des librairies en ligne de livres numériques**. Ce qui pose la question plus générale de la place des librairies dans le schéma recommandé.

3.5.2 La place des librairies dans le modèle économique d'*Europeana*

Il entre dans les objectifs explicites de la BnF pour *Europeana* le souhait d'inclure autant que faire se peut les libraires dans le modèle économique envisagé, objectif largement partagé par de nombreux éditeurs compte tenu du rôle bien évidemment décisif des libraires dans la chaîne du livre et des craintes que ceux-ci peuvent éprouver quant aux effets potentiels de la nouvelle économie numérique sur leurs métiers. En effet, d'une part les librairies en ligne « pure players », parfois filiales de grands groupes non européens, représentent une part croissante du commerce du livre ; d'autre part, la perspective de dématérialisation totale des livres représente a priori une menace encore plus forte sur l'activité traditionnelle de la librairie si elle se limite à exercer son métier de manière traditionnelle. C'est pourquoi de plus en plus de librairies, en commençant par les plus grandes, ont entrepris depuis plusieurs années de créer leur propre site de vente de livres en ligne. En France, un projet de « portail des librairies » permettant à de nombreuses librairies indépendantes de se regrouper pour disposer d'une offre de livres en ligne est par ailleurs en cours de réalisation sous l'égide du Syndicat de la Librairie Française. Enfin, certaines librairies autres que les « pure players » commencent à réfléchir à l'idée de vendre également sur leur(s) site(s) des livres en version numérique, en passant notamment des accords avec des diffuseurs numériques. En Grande-Bretagne, le cas du réseau de librairie WH Smith a été cité. En France, seule la librairie Lavoisier, pionnière, a jusqu'à présent expérimenté la vente de livres numériques... et, très marginalement, la FNAC. Mais d'autres devraient vraisemblablement les rejoindre dans les mois et années qui viennent, en France et dans le reste de l'Europe.

En tant que portail de recherche, *Europeana* sera un formidable outil d'information pour guider de nombreux types de publics vers les livres qui correspondent à leur recherche. Utilisant les technologies du numérique au profit de cette information du public, *Europeana* doit permettre également aux lecteurs de trouver où se procurer la version imprimée des livres qui les intéressent : un pourcentage extrêmement élevé de lecteurs devrait en effet continuer à préférer la lecture sur papier à la lecture sur écran pendant de

nombreuses années. Les librairies sont donc les partenaires normaux d'*Europeana* pour fournir dans ce cas les livres recherchés.

Par conséquent :

Notre recommandation est d'inclure les librairies dans le modèle économique d'*Europeana* à deux niveaux :

- 1. La fourniture directe des livres imprimés.**
- 2. L'accès payant aux livres numériques, par l'intermédiaire des opérateurs agréés, dès que suffisamment de librairies proposeront des sites de vente d'ebooks.**

Cette recommandation générale soulève les trois questions complémentaires suivantes :

- a. comment doit-on définir les librairies vers lesquelles seront effectuées des redirections ?
- b. en ce qui concerne les livres imprimés, comment l'interfaçage entre *Europeana* et les sites des librairies peut-il être effectué ?
- c. en ce qui concerne les livres numériques, quels devraient être les rôles respectifs des opérateurs agréés et des librairies ?

En reprenant ces trois questions dans l'ordre :

a. Il nous semble que la volonté de la BnF d'inclure les librairies dans le modèle économique envisagé n'est pas seulement un choix technique, mais également une décision politique de la bibliothèque en tant qu'institution publique destinée à préserver et à diffuser la culture européenne et le modèle ayant permis son épanouissement à travers les livres, qui repose sur la chaîne du livre. Face aux initiatives massives d'acteurs non européens « pure players » de l'Internet, qui pourraient sérieusement menacer ce modèle à de nombreux niveaux, il est légitime qu'*Europeana* puisse être un outil de promotion de cette conception européenne d'organisation de l'économie du livre.

Par conséquent, notre recommandation est de définir les librairies, au moins celles vers lesquelles *Europeana* proposera des liens pour l'achat de livres imprimés, comme des lieux physiques de vente de livres, d'accueil, de conseil et d'orientation du public.

Cette redirection ne pourra se faire qu'en faveur des librairies ayant déjà un site Internet, ou d'un portail de librairies, mais notre recommandation suppose que ces sites soient exclusivement ceux de librairies physiques traditionnelles – au moins pour la vente de livres imprimés.

b. En ce qui concerne les livres imprimés, une redirection vers un portail de librairie, au moins dans un cadre national, présenterait un avantage de simplicité comparable à celui de la redirection vers un opérateur agréé unique pour les contenus numériques. Ce point devra toutefois être validé pays par pays et en veillant à son acceptabilité par une collectivité très large de libraires et en fonction de la représentativité de ces éventuels portails. La définition des modalités d'interfaçage technique avec ces sites de librairies sort du strict cadre la présente étude. On peut toutefois faire les remarques suivantes :

- il est essentiel que les métadonnées utilisées par *Europeana* (voir plus bas) incluent une référence permettant d'identifier les ouvrages selon des normes standards internationales utilisées par les libraires. Les liens proposés par *Europeana* devraient en effet renvoyer, sur chaque site de librairie, vers la fiche produit précise du livre concerné, et non vers une page d'accueil globale, ce qui noierait la navigation et réduirait fortement l'intérêt de ces redirections.

- faute de redirection vers des portails nationaux, qui prendraient alors en charge ce point, il faudrait définir des modalités techniques de géolocalisation des internautes authentifiés, afin qu'ils puissent être orientés vers les sites des librairies les plus proches de chez eux, dans lesquelles ils pourraient finalement préférer se rendre pour feuilleter le livre et discuter avec le libraire avant de l'acheter. **Notre recommandation est également d'ajouter à la palette des outils de personnalisation prévus par Europeana une liste « Mes librairies » où l'internaute pourrait mémoriser ses librairies préférées.**

c. En ce qui concerne les livres numériques, **la redirection vers les sites de librairies vendant elles-mêmes des livres numériques devrait être gérée par les opérateurs agréés.** Mais compte tenu du très faible nombre de librairies, au sens où nous l'avons défini, qui vendent pour le moment des livres numériques, il ne serait pas raisonnable de proposer à ce stade qu'*Europeana* inclue obligatoirement et dès ses débuts, les libraires dans le circuit d'accès aux livres numériques. Toutefois, et afin que cet objectif ne reste pas un vœu pieux, **il est souhaitable d'inclure dans les accords contractuels avec les opérateurs de diffusion agréés une clause de « meilleur effort » pour associer des librairies à la vente des livres numériques, avec si possible un calendrier.** Cette association ne devrait cependant pas être conçue de manière trop exclusive ou rigide, afin de laisser jouer au fil du temps toute la souplesse des arrangements contractuels et du développement de ce nouveau marché. A titre de simple exemple, certains libraires pourraient se charger seulement de la vente par téléchargement, alors que des opérateurs agréés continueraient à gérer en direct les accès à une lecture en ligne sous la forme d'abonnements ou de pay-per-view.

3.5.3 Le cas particulier des lecteurs handicapés

Le numérique est une technologie extraordinaire au service des personnes handicapées pour faciliter leur accès à la lecture. En effet :

- les personnes mal voyantes peuvent utiliser les livres numériques pour adapter la taille des caractères à leur vue, l'édition numérique étant pour eux une forme nouvelle d'« édition adaptée » sur mesure ;
- les personnes non voyantes peuvent, avec certains formats de livres numériques comme le format Adobe PDF (si le fichier est en mode texte), utiliser des logiciels de synthèse vocale tels que le logiciel JAWS pour écouter les textes ou pour émuler des lecteurs Braille.
- les personnes handicapées motrices peuvent avoir un accès immédiat aux livres sans avoir besoin de se déplacer ; elles peuvent utiliser, lorsque l'ouvrage est téléchargé, une fonction de défilement automatique du texte dont la vitesse est réglable pour lire un ouvrage de manière autonome sans qu'on leur en tourne les pages, comme elles regarderaient un film ;

- certaines personnes handicapées mentales sont plus à l'aise avec la lecture de livres sur écran qu'avec la lecture de livres en papier, pour lesquels elles ont parfois des réactions de rejet.

Dans ce contexte, il est donc souhaitable qu'*Europeana* favorise au mieux les accès des handicapés à la lecture numérique.

Dans cet objectif, notre recommandation est de proposer aux lecteurs handicapés un lien vers les sites de bibliothèques numériques qui leur permettent un accès à distance et gratuit aux livres dans des conditions légales et contractuelles acceptées par les éditeurs.

En France, il existe en particulier depuis 2005 une bibliothèque publique qui remplit cette mission, à l'initiative de la ville de Boulogne Billancourt : **la Bibliothèque Numérique pour le Handicap**. Cette bibliothèque remplit toutes les conditions d'acceptabilité par les éditeurs car les livres qui y sont prêtés aux lecteurs handicapés sont achetés par la bibliothèque selon le modèle du nombre limité d'accès simultanés et font l'objet d'une protection par des systèmes DRM. Les lecteurs handicapés doivent préalablement s'inscrire, et passent la plupart du temps par une des associations de personnes handicapées qui soutiennent le projet. Le site est entièrement développé en respectant les normes d'accessibilité pour les handicapés visuels. Enfin, autre point important et différenciant ce projet de bien d'autres : tous les types de handicap sont concernés.

Bien entendu, cette redirection vers des bibliothèques numériques spécialisées au profit des lecteurs handicapés devrait pouvoir s'étendre à d'autres bibliothèques qui, en France ou dans d'autres pays européens, rempliraient ce même ensemble de conditions.

Faut-il étendre cette variante par rapport au modèle économique recommandé à d'autres sites de bibliothèques numériques en proposant depuis *Europeana* des liens vers toutes les bibliothèques numériques proposant un prêt à distance des ebooks à leurs lecteurs inscrits ? Nous ne le recommandons pas dans la phase actuelle, car cela risquerait de diminuer très fortement l'acceptabilité du projet par les éditeurs, et donc d'être un sérieux frein à son développement. Par ailleurs, cela impliquerait à nouveau une complexité trop forte dans la gestion et l'ergonomie du projet (liée notamment à l'identification en temps réel des livres disponibles dans chaque bibliothèque et au « bruit » causé par des informations non pertinentes pour les lecteurs non inscrits) qui serait un autre frein à sa mise en place opérationnelle dans les mois qui viennent. On peut enfin supposer que les lecteurs inscrits dans ces différentes bibliothèques y disposent des fonctionnalités de recherche pertinentes dans la base des livres disponibles sans avoir besoin de se connecter à *Europeana*. Ce point mériterait toutefois d'être rediscuté dans une phase ultérieure du projet avec les représentants des éditeurs et les bibliothécaires.



The screenshot shows a web browser window displaying the website for the 'Bibliothèque Numérique pour le Handicap'. At the top, there is a navigation bar with links: 'Aller au contenu', 'Plan du site | Mode d'emploi | Contacts'. To the right are logos for 'AFM', 'VILLE DE BOULOGNE BILLANCOURT', and 'BNH'. Below this is a menu with five items: 'Accueil', 'Catalogue', 'Avant d'emprunter', 'Réservations', and 'Recherche'. The main heading is 'Bibliothèque Numérique pour le Handicap'. Below the heading is the 'Elupha' logo, which features a stylized green leaf and white flowers. Underneath the logo is the text 'Elupha' and 'Partenaires Elupha'. The main content area contains several paragraphs of text describing the project's experimental phase, the ELUPHA study, and how to register. At the bottom of the page, there is a link: 'Site de la Bibliothèque Numérique pour le Handicap'.

3.5.4 L'accord des auteurs et des ayants droit

Les auteurs ou leurs ayants droit devraient bien évidemment être associés au déploiement des textes sous droits au sein d'*Europeana*, ce à plusieurs niveaux :

- étant donné l'ampleur et l'impact potentiel du projet, qui s'inscrit dès le départ partiellement en contrepoint d'initiatives respectant peu les titulaires de droits, les associations d'auteurs devraient être tenues régulièrement informées des avancées du projet et de ses principes de fonctionnement, et leur avis recueilli ;
- la commercialisation de versions numériques des ouvrages suppose **l'accord explicite des auteurs**, la cession de ce droit ne pouvant se présumer au titre du code de la propriété intellectuelle ;
- la simple indexation d'un ouvrage dans un moteur de recherche plein texte implique théoriquement une forme d'atteinte au droit moral imprescriptible des auteurs qui requiert également leur accord explicite. En effet, l'utilisation d'un moteur de recherche plein texte

et le parcours des extraits de livres correspondant à la recherche conduit à une modification de l'ordre de découverte de l'ouvrage prévu par son auteur.

Il n'appartient toutefois pas à *Europeana* de s'assurer elle-même de ces accords auprès des auteurs. Le processus normal de cette vérification doit suivre le parcours des accords contractuels :

- *Europeana* doit s'assurer par contrat que les opérateurs agréés « font leur affaire » de l'explicitation des droits d'auteurs auprès des éditeurs qu'ils représentent ;

- les opérateurs agréés doivent s'assurer de même par contrat que les éditeurs qu'ils diffusent « font leur affaire » de l'explicitation des droits auprès de leurs auteurs et ayants droit ;

C'est en définitive aux éditeurs seuls qu'il revient de prendre la responsabilité de proposer la diffusion d'un ouvrage sous forme numérique et de proposer son indexation dans un moteur de recherche en s'assurant d'un accord en bonne forme de leurs auteurs.

Cette recherche d'accord ne devrait poser aucun problème pour les nouveaux livres, les contrats des éditeurs souhaitant s'engager dans la diffusion numérique incluant désormais des clauses explicites à cet effet. Pour les ouvrages antérieurs, il appartient aux éditeurs de procéder aux vérifications ou explicitations qu'ils jugent nécessaires pour se conformer au droit. Pour les ouvrages anciens, la recherche des ayants droit et l'explicitation de leur accord peut s'avérer plus longue ; ce point devra être pris en compte pour l'éventuelle mise en place de plans de numérisation rétrospective. Dans cette perspective, la Société des Gens de Lettres propose, pour la France, de créer un **fichier centralisé des ayants droit** facilitant cette recherche. Un soutien à la création de fichiers de ce type, en France et en Europe, pourrait être envisagé dans le cadre du déploiement d'*Europeana*. Un intérêt corrélatif de tels fichiers pourrait être de constituer un outil, cette fois au service d'*Europeana*, pour vérifier autant que faire se peut la qualité d'« œuvres orphelines » de certains ouvrages dont elle pourrait prendre en charge la diffusion numérique, leurs éditeurs et leurs ayants droit n'ayant pas pu être identifiés. Des travaux de la commission européenne, autour de la création de mécanismes d'octroi de licences pour l'exploitation des œuvres orphelines - et des œuvres épuisées - sont en cours ; ils devraient prochainement faire l'objet de recommandations qui devront être considérées en vue du déploiement complet du projet *Europeana*, en collaboration avec les représentants des auteurs et des éditeurs.

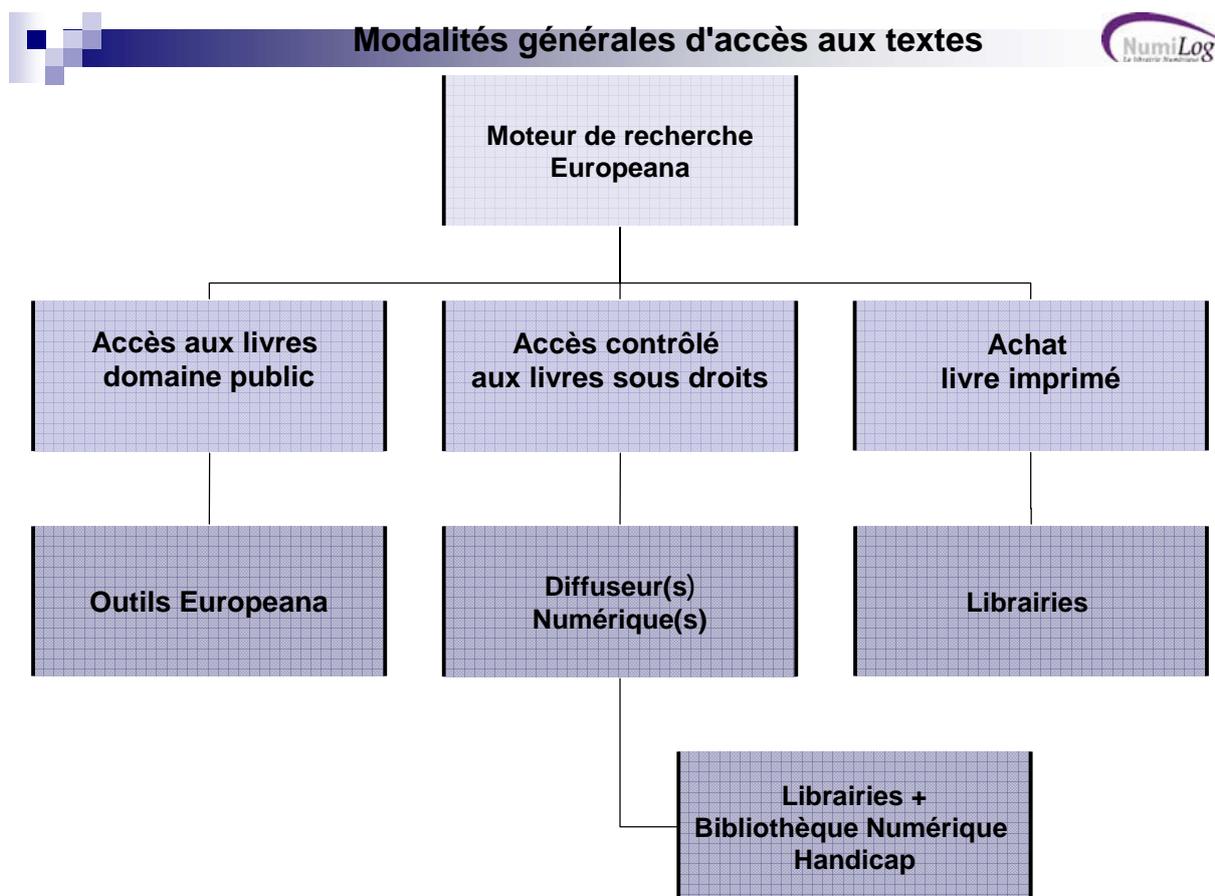
Enfin, certains auteurs ayant repris leurs droits sur leurs ouvrages dans le cadre prévu par le code de la propriété intellectuelle en cas d'arrêt prolongé d'exploitation commerciale, pourraient souhaiter apporter leurs ouvrages à *Europeana*. Il faudrait prévoir, dans une phase ultérieure, un dispositif facilitant cette possibilité, par exemple en indiquant aux auteurs une liste d'opérateurs agréés pouvant se charger de la diffusion numérique de leurs livres.

3.6 Le parcours type des usagers

On peut reformuler la recommandation proposée par les quatre principes suivants :

1. *Europeana* devrait fonctionner comme une bibliothèque numérique pour les livres du domaine public (et les œuvres orphelines) et comme **un portail de recherche pour les livres sous droits**.
2. En tant que portail de recherche, *Europeana* devrait proposer **une recherche gratuite commune dans les livres du domaine public et les livres sous droits**. Les critères de classement des livres par ordre de pertinence ne devraient évidemment pas tenir compte de cette différence, et la recherche devrait s'effectuer de manière homogène dans toute la base de livres indexés. Cela suppose en particulier que la structure des métadonnées associées aux livres soit la même pour tous les livres – et requiert donc un effort d'homogénéisation en ce sens.
3. Les modalités d'accès aux livres sous droits, numériques ou imprimés, devraient être organisées selon **un schéma d'« architecture distribuée »** entre différents opérateurs spécialisés : *Europeana*, diffuseurs numériques, libraires, bibliothèques pour le handicap.
4. Le modèle économique d'accès au texte intégral des livres sous droits en version numérique devrait être celui du **modèle de la librairie numérique**, avec un paiement par les usagers finaux selon des formules tarifaires variées.

Le schéma suivant résume, en même temps que le parcours d'un usager en fonction de la nature de sa recherche, les lieux de compétence de chaque acteur conformément à ces quatre principes.

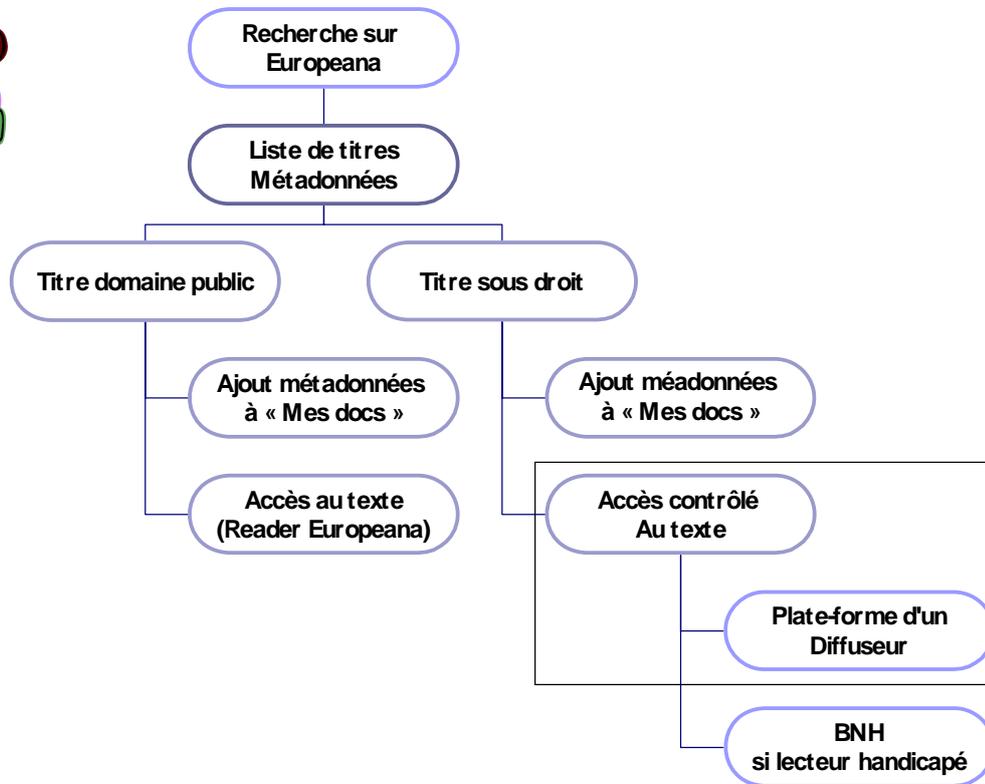


Le renvoi des diffuseurs numériques vers les librairies pourra avoir lieu dans une deuxième phase du projet, lorsque suffisamment de site de librairies en ligne auront développé des offres de vente de livres numériques, sous la forme de téléchargements à l'unité, voire également d'autres types de formules tarifaires. Dans une première phase, les diffuseurs, déjà opérationnels pour cela, pourront gérer totalement la vente des textes numériques aux utilisateurs finaux selon toutes les formules tarifaires envisagées. Le renvoi vers un accès gratuit pour les lecteurs handicapés est, comme signalé précédemment, d'ores et déjà possible en France grâce à la Bibliothèque Numérique pour le Handicap. Le renvoi vers les librairies pour l'achat des livres imprimés doit pouvoir être ajouté très rapidement, après des discussions sur les modalités précises avec les organisations représentant les libraires.

Le parcours type de navigation d'un usager sur le site *Europeana* peut alors être décrit plus en détail par les trois schémas suivants :

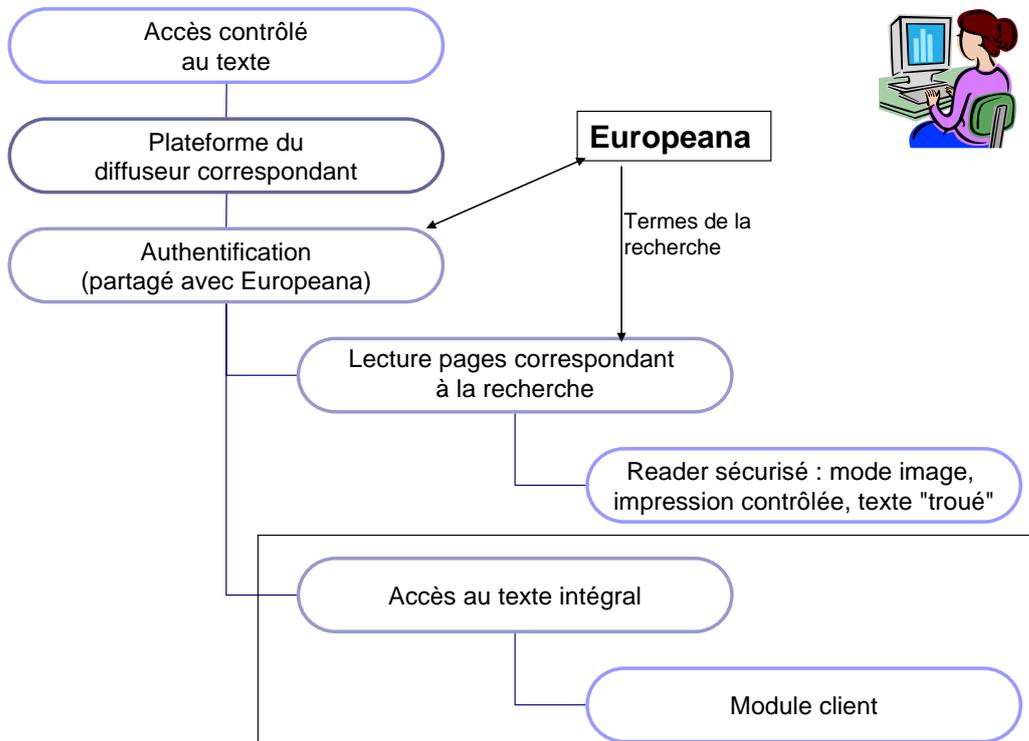


Le parcours de l'internaute (1)

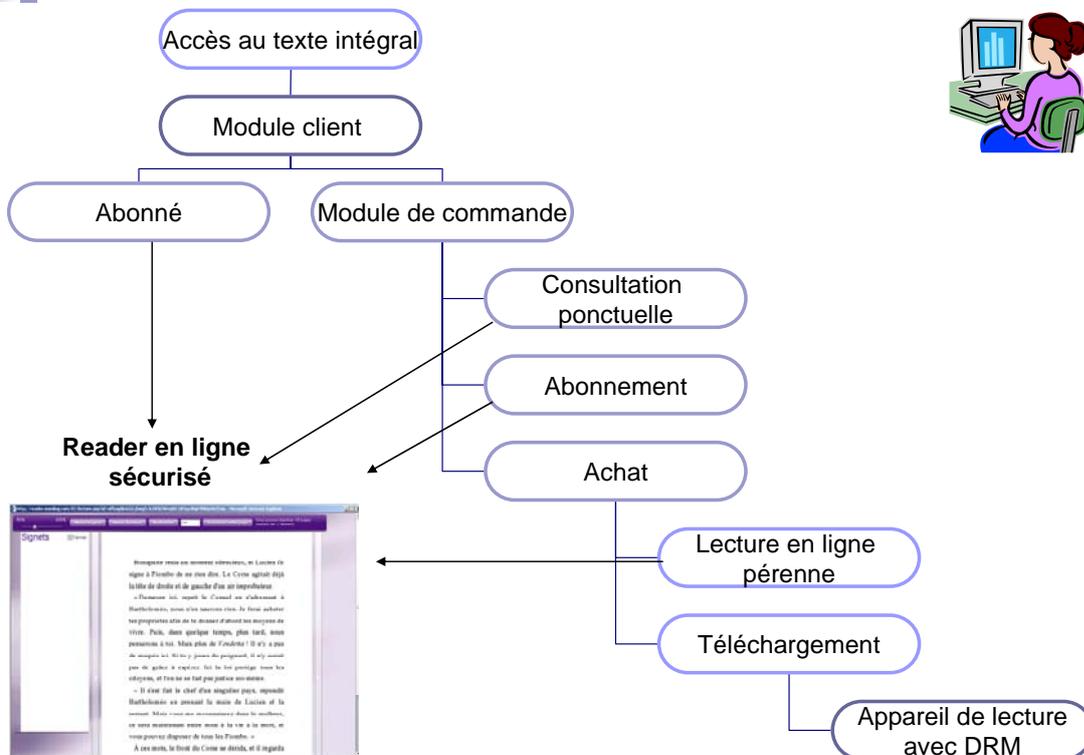




Le parcours de l'internaute (2)



Le parcours de l'internaute (3)



L'utilisation du moteur de recherche sur *Europeana* fournira à l'utilisateur une liste de titres dont il pourra ajouter les métadonnées à sa page « Mes Documents », qu'il s'agisse de titres du domaine public ou de titres sous droits. La visualisation de pages d'un des ouvrages correspondant à l'expression recherchée devrait se faire sur le site du diffuseur agréé avec un Reader équipé de solutions DRM ad hoc pour la lecture en ligne. En outre, même si la recherche se fait effectivement dans tout le texte, les pages offertes à la visualisation libre pourraient, selon les choix paramétrables des éditeurs, n'être que des extraits d'ouvrages, continus ou par exemple avec une page sur N (textes « troués »).

L'utilisateur devrait s'authentifier avant de pouvoir accéder à ces pages libres. **Il est important toutefois que ce parcours de visualisation demeure très fluide** et que l'accès aux dites pages sur le site du diffuseur numérique se fasse en « deux clics » au maximum, comme s'il s'agissait d'un service interne à *Europeana*.

Si l'internaute souhaite accéder au texte intégral, la plate-forme du diffuseur le dirigera vers un module client où il pourra soit être reconnu comme abonné, soit être orienté vers un module de commande où les différentes formules tarifaires lui seront présentées. A la place de ce module client unique, dans une phase ultérieure, la plate-forme du diffuseur pourra présenter à l'internaute une liste de sites de librairies numériques proposant cet ouvrage dans leurs catalogues.

3.7 Les contrats de partenariat avec les diffuseurs agréés et les transferts de données

Les opérateurs commerciaux chargés de l'accès contrôlé aux textes intégraux devraient être choisis sur une double base :

- choix par l'éditeur de son diffuseur au sein d'*Europeana*, à raison d'**un seul diffuseur par livre** dans le schéma *Europeana* initial envisagé ;
- agrément du diffuseur par *Europeana* sur la base de la signature d'un **contrat de partenariat**. Cette forme contractuelle permet en principe une souplesse juridique plus grande et correspond au principe du modèle recommandé : ces contrats ne visent pas à confier une prestation à des opérateurs privés ni à concéder une part de service public, mais bien à définir des coopérations bilatérales libres liées à la présentation et à l'accès à des contenus éditoriaux.

Le contrat type de partenariat devrait prévoir notamment :

- des engagements d'*Europeana* quant au **respect des droits de propriété intellectuelle**, aux conditions de sauvegarde et d'usage des données qui lui seront transférées, en garantissant en particulier leur usage pour la seule recherche et présentation des ouvrages ;
- des engagements d'*Europeana* quant au signalement et à la nécessaire **mise en valeur des ouvrages sous droits** sur son site (par exemple avec des logos appropriés les différenciant des ouvrages du domaine public) ;
- un **engagement juridique** des diffuseurs quant à la détention de contrats en bonne et due forme avec les éditeurs qu'ils prétendent représenter et dont ils apportent les contenus, les désignant notamment comme leur représentant au sein d'*Europeana* ;

- des **engagements déontologiques** des diffuseurs quant au signalement de leur partenariat avec *Europeana*, à leur démarche d'association des libraires à la vente de contenus numériques et éventuellement à leurs pratiques tarifaires (modèles de tarification variés et équilibrés – qu'il leur appartient de définir).
- des **engagements techniques** des diffuseurs portant sur la normalisation des données communiquées, sur les protocoles de communication et éventuellement sur quelques principes généraux.

Les paragraphes qui suivent décrivent à titre exploratoire quelques orientations recommandées pour le contenu des accords techniques avec les diffuseurs agréés.

3.7.1 Normalisation et transfert des métadonnées

Le référencement des ouvrages dans le moteur de recherche d'*Europeana* suppose que les diffuseurs agréés communiquent au serveur *Europeana* un ensemble des fichiers de métadonnées descriptives des livres pour la recherche et l'affichage des informations sur le livre. A cet égard, il est absolument nécessaire de **définir des normes homogènes** qui devront être respectées par tous les partenaires du projet. Pour cela, il faudra définir un compromis entre les normes déjà définies par la BnF pour l'actuel prototype d'*Europeana*, les pratiques et possibilités actuelles des diffuseurs numériques et des éditeurs, et les utilisations à prévoir des données, notamment par les libraires. Il est également nécessaire d'anticiper l'évolutivité de ces normes par enrichissement progressif.

Dans l'état actuel de nos informations, on peut identifier les premières hypothèses et recommandations suivantes :

a. Formats de métadonnées

Les métadonnées actuellement utilisées par la BnF sont au **format DUBLIN CORE**. **Nous recommandons toutefois que les opérateurs communiquent à *Europeana* des métadonnées au format ONIX** (Online Information eXchange), compatible avec la norme DUBLIN CORE. La communauté internationale du livre recommande en effet de plus en plus l'utilisation de la norme ONIX, au format XML, pour le maintien et la communication de données bibliographiques et commerciales entre les professionnels : éditeurs, diffuseurs, libraires. ONIX définit à la fois une liste de types de métadonnées et la manière de communiquer ces données dans un message ONIX. Cela permet de spécifier les types de données utilisés par un message (par exemple le type de classification thématique des ouvrages retenu), en s'assurant qu'il n'y a pas d'erreur d'interprétation, tout en laissant aux éditeurs une grande souplesse quant aux types de données renseignées. La norme ONIX est bien plus riche que la norme DUBLIN CORE, qui contient 15 champs de données possibles, car elle contient beaucoup plus d'éléments de base qu'il est possible de renseigner.

Les types de données qui devraient être communiquées sous un format ONIX par les opérateurs agréés sont au minimum celles qui sont renseignées pour les ouvrages présentés sur le prototype *Europeana*, soit :

- **Titre**
- **Auteur(s)**
- **Editeur**
- **Date de publication**
- **Classification (ou « sujet »)** : le type de classification peut être choisi par l'éditeur, mais une normalisation est à envisager – par exemple la norme DEWEY.
- **Type** : soit, pour les livres, monographie imprimée, en un ou plusieurs volumes.
- **Langue**
- **Format numérique** : on peut indiquer ici les multiples formats sous lesquels le livre numérique est disponible, par exemple PDF et PRC.
- **Droits** : ouvrage sous droits.
- **Identifiant** : l'identifiant recommandé est a priori l'ISBN-13 (après conversion si nécessaire), qui devrait permettre un lien avec les catalogues des libraires.
- **Provenance** : l'opérateur de diffusion qui apporte cet ouvrage à *Europeana* – information qui permettra de faire le lien vers le site de cet opérateur pour accéder au contenu du livre.

Par ailleurs, il devrait être possible de proposer à *Europeana* d'offrir aux usagers des métadonnées supplémentaires si elles sont fournies par l'opérateur de diffusion. En particulier, il paraît souhaitable, en fonction de la nature des livres, d'ajouter les données suivantes, **très importantes pour la présentation et la mise en valeur des livres** et qu'on peut inclure dans un message ONIX :

- **Table des matières**
- **Couverture**
- **Descriptif ou 4^e de couverture**

b. Transfert des métadonnées

Afin de répondre aux protocoles retenus par la BnF, les métadonnées devraient être mises à sa disposition sur des **serveurs OAI** fournis par les diffuseurs agréés. Ce type de serveurs permet de moissonner un entrepôt OAI, c'est-à-dire une base de données qui supporte le protocole OAI-PMH. Mais ce protocole ne permet pas directement de moissonner des données au format ONIX. On peut alors identifier trois hypothèses :

H1. Les métadonnées au format ONIX sont encapsulées dans des fichiers DUBLIN CORE par les diffuseurs, ce qui est une procédure non standard mais possible en utilisant l'un des champs du DUBLIN CORE, le champ « descriptif », pour y placer un lien vers une adresse Internet où se trouvera le message ONIX. Il appartiendra alors à *Europeana* de développer les procédures informatiques d'exploitation des champs de ces messages ONIX.

H2. On renonce à fournir des métadonnées au format ONIX et les diffuseurs s'engagent à convertir leurs métadonnées source au format DUBLIN CORE. L'inconvénient de cette solution est qu'il faut alors soit renoncer aux types de métadonnées que le DUBLIN CORE

ne permet pas de décrire, soit prévoir une procédure complémentaire de transfert, différente de la procédure OAI, pour communiquer ces métadonnées complémentaires (avec les fichiers texte par exemple).

H3. Il est possible de développer un protocole d'échange de données différent de l'OAI mais assurant les mêmes fonctions afin de récupérer automatiquement des métadonnées au format ONIX sur des serveurs destinés à cela. *Europeana* devrait dans ce cas programmer par ailleurs une requête périodique de ces serveurs selon ce protocole, ce qui est tout à fait faisable techniquement.

c. Remarques complémentaires :

La question de l'identifiant des livres soulève plusieurs types de questions complémentaires, qui méritent d'être approfondies au fur et à mesure du déploiement du projet :

- une question d'ordre « ontologique » : quelle est l'unité élémentaire de documentation que devrait référencer *Europeana* ? L'ISBN, recommandé en tant qu'identifiant actuellement le plus en usage, suppose que l'entité de base est une « publication », c'est-à-dire une œuvre éditée sous une certaine forme à une certaine date. Mais il serait envisageable également de s'attacher, notamment grâce à la forme numérique, à des « grains » plus fins, comme les chapitres. A contrario, on pourrait considérer, toujours en raison de la forme numérique, que l'œuvre représente le – ou au moins un – niveau pertinent de classement si ce sont les textes de ces œuvres qui intéressent les usagers. **Il faudrait alors recommander d'utiliser comme identifiant complémentaire le code ISTC**, proposé pour identifier toutes les « œuvres textuelles ».

- quelques éditeurs ou diffuseurs devraient être à même de **fournir un identifiant DOI** (Digital Object Identifier). Le DOI est un identifiant utilisé d'une manière générale pour différents types de ressources numériques en ligne. Il présente l'avantage d'être plus souple et plus riche que l'ISBN, en permettant notamment d'indiquer des droits sur l'objet numérique et une adresse URL où il se trouve. Il est utilisé par les éditeurs scientifiques qui souhaitent que leurs livres ou revues fassent partie du système de référencement bibliographique croisé CrossRef. *Europeana* pourrait, dès le début ou dans une deuxième phase, utiliser le DOI au moins comme deuxième identifiant quand il est fourni par l'opérateur de diffusion. Son utilisation obligatoire n'est toutefois pas recommandée à ce stade, pas plus que celle du système voisin ARK utilisé par la BnF, en raison de son encore très faible usage chez les éditeurs comme chez les libraires.

- la référence à l'ISBN comme identifiant du livre suppose que ce qui est référencé par *Europeana*, ce sont des livres imprimés, dont on fournit par ailleurs un accès à des versions numérisées. Mais, comme le font désormais la plupart des éditeurs scientifiques anglo-saxons, les versions numériques devraient en toute logique avoir leur propre ISBN car il s'agit de nouvelles publications. Faudrait-il alors **retenir comme identifiant principal sur Europeana l'ISBN du livre imprimé ou celui du livre numérique ? Devrait-on au minimum ajouter aux métadonnées une référence aux ISBN des versions numériques ?**

3.7.2 Normalisation et transfert des fichiers texte

Le référencement des ouvrages dans le moteur de recherche d'*Europeana* suppose par ailleurs que les diffuseurs agréés fournissent à ce moteur un accès au texte intégral afin d'y effectuer ses recherches.

Là encore, dans l'état actuel de nos informations, on peut identifier les 3 hypothèses suivantes :

J1. Les fichiers textes actuellement utilisés par la BnF pour l'indexation des livres dans le moteur de recherche *Europeana* sont au **format XML Alto**. Ces fichiers permettent au moteur de recherche de localiser et de surligner les différents endroits où un terme recherché a été trouvé dans le livre. L'avantage du format XML Alto est d'être fourni automatiquement par les prestataires de numérisation identifiée par la BnF pour les ouvrages du domaine public et qui pourraient également se charger de la numérisation massive d'ouvrages sous droits dont les sources sont en papier. La première hypothèse est donc de recommander aux diffuseurs de fournir les fichiers texte des ouvrages sous cette même norme. Un autre avantage est que ces fichiers au format XML ne sont en aucun cas des fichiers qui pourraient faire l'objet de présentation au public pour une lecture, ce qui peut, si cela s'avérait nécessaire, apporter une garantie supplémentaire aux éditeurs quant à l'usage limité à la recherche qui peut en être fait. Par contre, une réflexion sur le caractère approprié de cette norme pour l'indexation d'ouvrages européens dans diverses langues et graphies mériterait d'être poursuivie. Par ailleurs, l'inconvénient pratique de cette norme pour les livres récents est qu'elle impliquerait une conversion spécifique, et donc un coût supplémentaire pour les éditeurs, car ce format n'est pas fourni par la chaîne de production standard des livres, imprimés ou numériques.

J2. A titre alternatif, et au moins pour les ouvrages récents, les éditeurs et les diffuseurs qui le souhaitent devraient pouvoir communiquer à *Europeana* des fichiers au **format PDF en mode texte**. L'avantage du format PDF texte est de correspondre à un standard plus répandu d'archivage numérique par les éditeurs pour des ouvrages récents pour lesquels ils disposent généralement de fichiers issus de leur chaîne de production, sans nécessité de nouvelle adaptation. Cela facilitera une livraison plus rapide et moins coûteuse des fichiers de ces ouvrages. Cela suppose par contre qu'*Europeana* prenne en charge le développement des programmes permettant la recherche plein texte à partir de ce type de fichiers.

J3. Certains éditeurs pourraient avoir, malgré les garanties apportées, une réticence à livrer à *Europeana* tout type de fichier contenant le texte intégral des livres, en voulant s'assurer la maîtrise intégrale de ces fichiers. S'il s'agit d'éditeurs disposant de fonds importants et pertinents, il serait dommage de priver *Europeana* de leur collaboration sur ce motif. Une solution alternative peut alors être recommandée : l'utilisation de la technologie **OPEN SEARCH™** (développée par Amazon sous une licence *Creative Commons*). Cette technologie permet à un moteur « agrégateur » (par exemple A9.com dans le schéma ci-dessous) de collecter les résultats de recherche plein texte effectués par d'autres moteurs (par exemple Icerocket dans le schéma ci-dessous) :



Protocole OpenSearch™

En utilisant la technologie Open Search™, *Europeana* pourrait donc agréger à ses propres résultats de recherche les résultats de moteurs effectuant des recherches dans des bases d'indexation qui resteraient hébergées par les éditeurs ou leurs diffuseurs. Cela impliquerait, sur le plan technique, d'une part qu'*Europeana* développe un « méta moteur » effectuant cette agrégation, d'autre part que les éditeurs ou les diffuseurs s'équipent en serveurs assez puissants pour être l'objet des requêtes de ce méta moteur de manière fluide.

Nous recommandons que l'ensemble des hypothèses de normalisation et de transfert H1 à H3 et J1 à J3 fasse l'objet d'une validation et d'une spécification plus détaillée avec la BnF et un groupe de professionnels, diffuseurs, éditeurs et libraires, avant de démarrer les opérations de référencement des ouvrages sous droits.

Cette normalisation est en tout cas tout à fait essentielle pour le bon fonctionnement et le développement du projet dans des conditions efficaces et extensibles.

3.7.3 Protocole d'authentification partagé

Sur la plate-forme des opérateurs de diffusion, sera proposé un accès en ligne contrôlé mais gratuit à quelques pages des ouvrages correspondant à la recherche effectuée par l'utilisateur. Par ailleurs, l'internaute devrait pouvoir, pendant ou après cette recherche, ajouter les métadonnées de l'ouvrage parcouru à sa page « Mes Documents » sur *Europeana*, afin de constituer une liste organisée de favoris qui pourra comprendre des ouvrages provenant de fonds de livres du domaine public ou de livres sous droits diffusés par différents opérateurs. Enfin, différentes formules de paiement seront proposées dans un espace client pour accéder aux textes intégraux, et notamment des formules d'abonnement qui impliquent de reconnaître l'utilisateur.

Toutes ces opérations – et au minimum les deux premières – requièrent **une authentification unique des usagers partagée avec *Europeana*.**

A cette fin, nous recommandons d'utiliser un système d'authentification centralisé par un système de tickets selon le système « CAS » (pour [Central Authentication Service](#)). Le premier intérêt de ce système est de permettre de dédier la responsabilité de l'authentification à un seul acteur. Dans le système CAS, l'authentification se fait en garantissant la confidentialité des identifiants et mots de passe et en évitant toute nécessité de transmission et de mise à jour de bases d'identifiants et de mots de passe. En effet, les sites où s'effectuent l'authentification (en l'occurrence ceux des opérateurs de diffusion) se contentent de vérifier, non pas directement la validité des identifiants et mots de passe, mais celle de « tickets d'entrée » délivrés par le serveur CAS (en l'occurrence celui d'*Europeana*) aux usagers.

La figure suivante montre le chemin d'écoulement d'une authentification d'un système ayant un serveur CAS intégré.

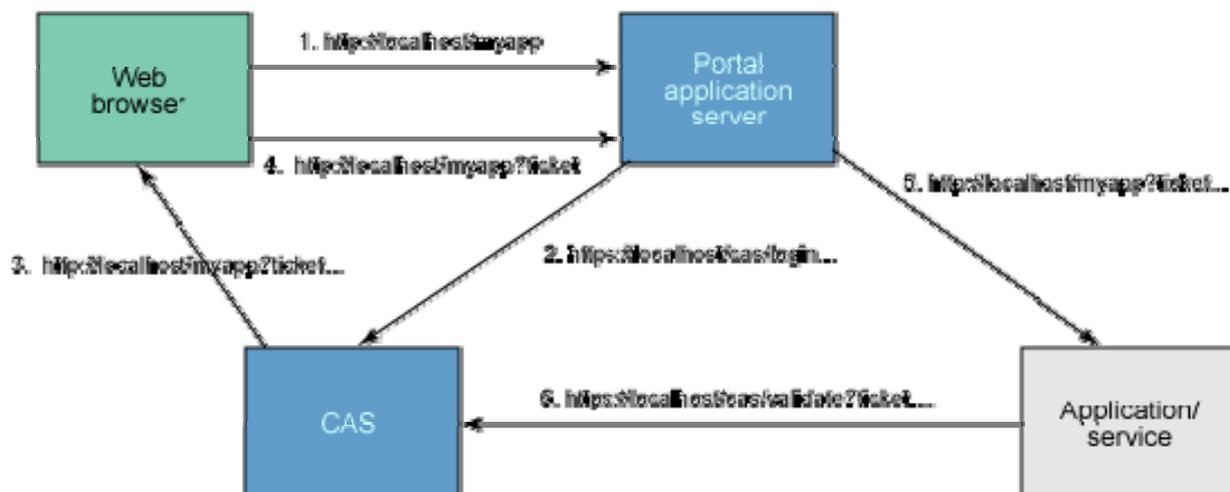


Schéma d'authentification par un protocole CAS

Voici les principales étapes dans le processus d'authentification :

1. L'utilisateur essaie d'accéder à une application en utilisant son URL. Il est réorienté sur la page d'authentification du service CAS par l'intermédiaire d'une connexion sécurisée HTTPS, transmettant le nom du service comme paramètre. L'utilisateur est invité à s'authentifier à l'aide d'une boîte de dialogue classique login/mot de passe.
2. L'utilisateur s'authentifie sur le serveur CAS.
3. Le serveur CAS réoriente l'utilisateur vers l'application cible, ajoutant un paramètre appelé *ticket* à l'URL. Le serveur CAS crée un ticket qui sera ensuite utilisé chaque fois qu'une authentification sera nécessaire.
4. L'application valide le ticket auprès du serveur CAS puis autorise, ou non, l'utilisateur d'accéder au contenu restreint.

Enfin, les avantages techniques de cette technologie sont multiples :

- la solution serveur est facile à installer et peut être développée dans un grand nombre de langages informatiques : Java, ASP, Perl, PHP... ;
- elle est facile à implémenter pour les sites utilisateurs ;
- elle est compatible avec les annuaires LDAP.

4 La numérisation des contenus et la politique documentaire

4.1 Processus et coûts de numérisation

Sur le plan technique, les modalités possibles de numérisation dépendent :

- d'une part des **sources disponibles** pour les livres ;
- d'autre part des **objectifs visés** pour la qualité et le format des livres numérisés.

Bien entendu, les différentes modalités de numérisation possibles impliquent des coûts de production différents, qui ont un impact économique sur les politiques de numérisation possibles.

En ce qui concerne les sources, on peut distinguer les deux principaux cas suivants :

4.1.1 Numérisation stricto sensu

Dans le cas de **livres de plus de 5 à 7 ans**, les éditeurs ne disposent le plus souvent que de sources papier. Dans ce cas, il faut procéder à une numérisation stricto sensu, parfois appelée **retro-conversion** des fonds éditoriaux. Un cas particulier au sein de cette catégorie est représenté par les **ouvrages indisponibles**, dont l'éditeur ne dispose même plus de source imprimée. Le processus de numérisation est alors le même, mais il faudra ajouter à ses coûts celui de la recherche de l'ouvrage, notamment dans les collections des bibliothèques, ou de son achat sur le marché des livres d'occasion.

La numérisation stricto sensu peut être confiée à des prestataires qui traitent des quantités massives de livres imprimés avec de nouveaux types de machines adaptées à cette numérisation massive pour des coûts unitaires relativement faibles. Alors que les coûts de numérisation sont de l'ordre de 0,25 € à 0,30 € la page si on procède à des numérisations en petite quantité, ils sont dans le cas de numérisation massive par ce type de prestataire d'ouvrages massicotables plutôt de l'ordre de 0,05 € à 0,06 €. **Un ouvrage de 250 pages peut donc être numérisé pour environ 12 à 15€**

La numérisation effectuée est en outre une numérisation intelligente (et qui tend à l'être de plus en plus), qui fournit au moins les types de fichiers suivants :

- fichiers images de toutes les pages de l'ouvrage, par exemple au format TIFF ou PNG ;
- fichier texte issu d'un processus de reconnaissance optique, qui peut être normé pour être supérieur ou égal à 95% ou à 98%, par exemple au format TXT ;
- fichier d'indexation, par exemple au format XML Alto, en vue d'une recherche plein texte.

4.1.2 Conversion ou adaptation numérique

En ce qui concerne **les nouveautés et les livres récents**, les éditeurs disposent en général de sources numériques, issues de leur chaîne de production, interne ou sous-traitée. En moyenne, c'est le cas chez de nombreux éditeurs depuis environ 5 à 7 ans (années 2000-2002). Dans ce cas, plutôt que de parler stricto sensu de « numérisation », il faut parler de **conversion** ou d'**adaptation numérique**, les fichiers source utilisés pour la PAO étant à des formats qui ne conviennent pas pour les besoins de la diffusion et de la lecture numériques.

L'adaptation numérique peut être réalisée par l'éditeur lui-même, par ses prestataires habituels (ceux qui mettent en page ses ouvrages), ou par des prestataires spécialisés : plateformes d'archivage, diffuseurs numériques, sociétés spécialisées dans la numérisation. Dans tous les cas, le processus d'adaptation et le coût dépendent largement à nouveau des sources et des objectifs. On peut distinguer les principaux cas suivants :

- la source peut être un fichier PAO, par exemple au format Quark Xpress™ ou Adobe InDesign™, ou un fichier au format Adobe PDF préparé pour l'impression ;
- l'objectif peut être de réaliser un fichier au format Adobe PDF enrichi et optimisé pour la lecture à l'écran et le téléchargement, ou un fichier à l'un des formats, Mobipocket PRC, Microsoft LIT, Palm PDB, plus adaptés à la lecture après téléchargement sur des supports mobiles.

Il est difficile de fournir un coût général pour ces prestations, car celles-ci dépendent largement de ces différents cas de figure ; la taille du livre, la complexité de la mise en page, la présence ou non d'illustrations, les types de polices de caractères utilisées et les enrichissements désirés pouvant également entrer en ligne de compte pour le calcul de ce coût. A simple titre de point de repère, **le coût moyen de réalisation d'un livre numérique optimisé est environ de 80€ à partir de sources PAO et de 60€ à partir d'un PDF « imprimeur »** pour un ouvrage de 250 pages sans spécificités de polices ou d'illustrations particulières.

S'il s'agit de réaliser des livres numériques à des formats comme le Mobipocket PRC, les coûts peuvent être du même ordre... ou bien plus élevés si la mise en page initiale, complexe, requiert une adaptation particulière pour une lecture sur des écrans d'agendas électroniques (par exemple la réalisation d'un moteur de recherche spécifique pour un dictionnaire).

A ces coûts généraux de production, il faut parfois ajouter des coûts dits de « désarchivage », qui peuvent atteindre 50€ par ouvrage, réclamés par les prestataires de composition de certains éditeurs pour leur fournir les fichiers source dont les éditeurs ne disposent pas sur leurs propres systèmes informatiques.

4.1.3 Avantages comparatifs

Contrairement à une idée répandue, **le coût de « numérisation » au sens large du terme est donc beaucoup plus élevé quand l'ouvrage est déjà disponible sous une forme numérique** que quand il faut le scanner. Ce fait n'est pas un paradoxe : il est simplement dû au fait que les objectifs visés pour les fichiers finaux ne sont pas du tout les mêmes et que les tâches résiduelles d'adaptation ne sont pas toutes automatisables... ou pas encore toutes automatisées.

Par exemple, dans le cadre d'une adaptation numérique, tout ou partie des tâches suivantes devra être effectué pour un livre au format PDF :

- Vérification de l'incorporation correcte de toutes les polices de caractères
- Détournement et suppression des traits de coupe
- Réalisation d'une table des matières hypertexte
- Création de signets de navigation affichés à l'ouverture
- Activation de tous les liens web du texte
- Réalisation et « épingle » d'une mini-couverture
- Optimisation afin de garantir un équilibre poids/qualité en vue d'une lecture écran et/ou d'une impression partielle de pages

Ces processus d'adaptation seront probablement facilités à l'avenir et les coûts réduits :

- par la mise en place au sein des chaînes de production des éditeurs ou de leurs prestataires de flux de production de fichiers structurés en XML à partir desquels la production des fichiers de livres numériques à des formats variés associés à des feuilles de styles spécifiques sera plus largement automatisée (sans toutefois être totale) ;
- par le recours à des plates-formes spécialisées dans l'archivage et la numérisation des fichiers mettant en œuvre des procédés de conversion standardisés, qui dégageront des économies d'échelle.

Mais les décisions de numérisation qu'auront à prendre les éditeurs aujourd'hui et dans les prochains mois reposeront le plus souvent sur les ordres de grandeur économiques qui viennent d'être présentés.

Au-delà des coûts, nettement à l'avantage de la numérisation à partir du papier si elle concerne des quantités importantes d'ouvrages, **les résultats obtenus sont nettement à l'avantage de la conversion de sources numériques en mode texte :**

- La numérisation stricto sensu fournit des fichiers en mode image (et des fichiers texte destinés seulement à la recherche), qui sont, on l'a dit, de meilleurs garants de la protection des ouvrages lors de leur lecture en ligne. Mais rien n'empêche, à partir de fichiers en mode texte, de réaliser des fichiers en mode image pour la seule lecture en ligne, selon des processus automatisés peu coûteux.
- Par contre, l'adaptation numérique à partir de fichiers PAO fournit des fichiers en mode texte présentant de très nombreux avantages pour la lecture hors ligne :
 - Même si on obtient en mode image des niveaux de reconnaissance optique avec des qualités de 98%, les fichiers obtenus ne sont jamais parfaits. La recherche plein

texte contient toujours des erreurs résiduelles ou des manques. Et ces ouvrages peuvent moins bien servir les logiciels de synthèse vocale pour les personnes non voyantes.

- La qualité de lecture à l'écran d'un fichier en mode texte à tout niveau de zoom est en général très supérieure à celle d'un fichier image.
- Le poids des fichiers en mode texte est en général bien inférieur à celui des fichiers en mode image.
- Les fichiers en mode image ne sont pas adaptés du tout à la lecture sur les écrans des supports mobiles de type PDA ou des lecteurs à base d'encre électronique. Ceux-ci supposent en effet d'utiliser des techniques de « reflow », qui recalculent la pagination du texte en fonction de la taille de l'écran et de la taille des caractères choisie par le lecteur – ce qui est impossible en mode image. Le passage des fichiers image à des fichiers texte préparés pour une conversion à ces formats serait imparfait, car d'une part la reconnaissance optique n'atteint jamais une qualité de 100%, d'autre part une remise en page s'avérerait nécessaire.
- Les sources PAO sont plus compatibles avec d'éventuels enrichissements destinés à différencier les versions numériques des versions papier, comme par exemple une re-mise en page spécifique adaptée à l'écran, l'ajout de fichiers son, de vidéo, ou de compléments d'information appelables sous forme de « pop up » dans le texte ;
- Enfin, si on souhaite utiliser les fichiers numérisés pour une impression à la demande, les fichiers en mode texte sont en principe de meilleure qualité que les fichiers en mode image – mais ce point est à valider avec les prestataires de numérisation.

Il faut donc, selon les types de sources disponibles et les objectifs visés, tenir compte de ces deux types de paramètres pour définir une politique de numérisation adaptée. Notre recommandation générale est la suivante :

Il semble souhaitable de procéder à une numérisation stricto sensu en mode image avec un traitement OCR des sources papier lorsque seules celles-ci sont disponibles, mais de privilégier une conversion des sources numériques en mode texte chaque fois que cela est possible, malgré ses coûts plus élevés.

Cette recommandation générale est sans doute à relativiser en fonction des usages et des objectifs commerciaux que les éditeurs associeront à la numérisation de chaque ouvrage.

4.2 Choix de numérisation et politique documentaire

A priori, la politique de numérisation des contenus sous droits est de la seule responsabilité des éditeurs. Toutefois, l'ampleur publique et européenne d'*Europeana*, sa visibilité potentielle importante et la taille du catalogue de titres qu'elle est susceptible de réunir peuvent être de puissants facteurs d'incitation des éditeurs à numériser leurs contenus afin de bénéficier de cette visibilité. De plus, la Bibliothèque numérique européenne pourrait jouer un rôle d'aide technique et indirectement d'aide financière dans la numérisation de certains

fonds, en particulier par l'intermédiaire d'accords quant à l'usage de fonds destinés à cette numérisation et gérés par le Centre National du Livre.

Les **choix de numérisation prioritaires des éditeurs** prennent et prendront sans doute en compte les paramètres hétérogènes suivants :

- **Les cibles de lectorat et les usages envisagés** afin de rentabiliser la numérisation, selon la nature des livres. Si on regarde les débuts du marché des livres numériques, les trois types de couples usage/catégorie de livres qui fonctionnent commercialement le mieux sont les suivants :

- Les essais, ouvrages universitaires ou techniques, consultés sur ordinateur, à destination des étudiants, enseignants ou professionnels concernés, en particulier les ouvrages de référence, correspondant à l'actualité ou aux programmes d'enseignement en cours ;
- Les ouvrages de type guides de voyage ou dictionnaires, linguistiques ou professionnels, consultés sur des supports mobiles ;
- Les livres de fiction récents, versions numériques des « best-sellers » de librairie associés à des auteurs renommés, lus sur divers types de supports.

On peut supposer que dans un avenir proche les livres pratiques et les bandes dessinées auront également un succès numérique important. Les formats de numérisation retenus et la nature des enrichissements adoptés afin de distinguer les versions numériques des versions papier dépendent de ces usages envisagés. Les modèles commerciaux privilégiés également.

- Les **risques économiques** associés par certains éditeurs à une éventuelle cannibalisation de la vente de certains ouvrages « vedettes », qui représentent une part importante de leur chiffre d'affaires par la diffusion de versions numériques.

- La volonté d'**utiliser le numérique pour modifier le cycle de vie de certains livres**, en continuant de les exploiter au-delà de la période habituellement retenue pour leur exploitation sous forme papier. Toutes les nouveautés produites en version numérique auront par nature désormais une durée de vie commerciale plus longue. La question de la numérisation rétrospective d'ouvrages dits « épuisés » se posera, en particulier pour les éditeurs de sciences humaines et sociales.

- Des **choix techniques** liés à la numérisation des ouvrages pour lesquels on ne dispose que de sources papier. Au-delà des choix stratégiques ou de la politique documentaire, en effet, tous les ouvrages ne sont pas égaux devant la numérisation selon le type de papier utilisé ou la typographie. Il faut sans aucun doute effectuer ici d'abord des choix pratiques pour maximiser la qualité des numérisations effectuées.

- Des **choix juridiques** liés à la difficulté plus ou moins grande de réunir dans des conditions claires l'accord des ayants droit. Les ouvrages comportant notamment une grande richesse iconographique, quels que soient leur valeur documentaire et l'intérêt d'en proposer des versions numériques, poseront plus de difficultés pour clarifier cet accord.

Dans ce contexte, une **politique documentaire d'Europeana peut représenter pour certains éditeurs un guide important pour leurs choix de numérisation.**

Bien entendu, les éditeurs souhaiteront adopter une politique de numérisation leur permettant d'envisager une rentabilisation de cette numérisation dans le cadre de canaux de distribution divers. En ce qui concerne les nouveautés, la priorité sera accordée aux indications fournies par le marché, liées par ailleurs pour le moment en grande partie aux achats de bibliothèques. Et il est souhaitable qu'Europeana puisse accueillir tous ou la plupart des ouvrages ainsi sélectionnés par les éditeurs, sans discrimination thématique, au-delà de certains principes déontologiques, afin de devenir un portail de recherche concurrentiel avec les portails privés et à destination du plus grand nombre de lecteurs. Mais d'autres ouvrages, plus anciens ou au cycle de vente plus long, ne feraient pas l'objet d'une numérisation rapide sans une incitation des pouvoirs publics sous la forme d'aides économiques et du bénéfice des bas coûts de numérisation permis par les numérisations massives. Or ces ouvrages peuvent avoir un intérêt significatif pour des usages d'étude, et d'une manière générale pour le projet de constituer une base d'ouvrages représentatifs de la culture européenne et de son histoire. Il semble approprié, pour ce type d'ouvrages, qu'Europeana indique aux éditeurs quels sont ses choix documentaires prioritaires, en contrepartie des éventuelles aides apportées.

Notre recommandation est donc qu'Europeana, tout en demeurant ouverte aux ouvrages numérisés librement par les éditeurs, rende publique une politique documentaire assez précise, qui servirait de « guide line » et de cahier des charges pour la négociation de collaborations techniques et financières en vue de la numérisation d'ouvrages correspondant à cette politique.

Cette **stratégie mixte** représente un équilibre, qui nous semble acceptable par la plupart des éditeurs, entre une orientation volontariste donnée à une politique menée en partie sur fonds publics dans un objectif d'intérêt général, et un partenariat ouvert et sans discrimination catégorielle, favorable à la croissance du fonds d'ouvrages numériques référencés. Cette suggestion devra toutefois faire l'objet de validations complémentaires.

Il n'est pas dans l'objet de la présente étude d'entrer dans le détail d'une telle politique documentaire. De premiers axes ont été identifiés par la BnF pour les contenus en français, dans le prolongement et l'enrichissement la politique documentaire de *Gallica*. Ils privilégient les ouvrages concernant le patrimoine national (critique littéraire, histoire de France, régions françaises, œuvres des scientifiques français, dictionnaires...), des ouvrages à vocation internationale (apprentissage du français, traduction de chefs-d'œuvre d'autres pays...) et des ouvrages concernant l'Europe en général (identités nationales, sciences et techniques, questions économiques et sociales, relations internationales...). Si l'idée d'une politique documentaire est retenue, ces axes devront être détaillés, en collaboration avec des éditeurs, jusqu'à identifier les ouvrages ou collections d'ouvrages susceptibles d'être concernés.

4.3 Simulation économique

Plusieurs milliers (approximativement 15 000 à 20 000) de livres numériques en français sont déjà proposés sur le marché et pourraient déjà, si leurs éditeurs le souhaitent, être référencés dans *Europeana*. Pour ces livres, la question de la rentabilisation potentielle des coûts de numérisation a déjà été traitée. Pour les centaines de milliers d'autres numérisables, cette question se pose dans le cadre du modèle économique proposé afin de vérifier que ce modèle permettrait à *Europeana* d'espérer dans les prochaines années qu'un volume important de livres numériques soit ajouté et proposé aux internautes.

Compte tenu du modèle économique recommandé :

- Aucun coût variable direct n'est à prévoir par la BnF pour développer le fonds de livres sous droits qui sera présenté dans *Europeana*. Les seuls coûts à prévoir sont les éventuels coûts de développement des solutions informatiques qui seront retenues, en accord avec certaines de nos propositions, pour le fonctionnement du système et en particulier pour la mise au point de protocoles d'échanges de données avec les diffuseurs agréés. Ces coûts, qui doivent faire l'objet d'analyses complémentaires, ne sont pas considérés ici, mais devraient représenter un enjeu économique relativement bas eu égard à l'ensemble de l'économie du projet.

- Il appartient aux éditeurs d'effectuer leurs choix de numérisation – en prenant les risques économiques associés aux coûts représentés par cette numérisation. Il est toutefois intéressant de **vérifier que de tels choix seront envisageables avec des hypothèses raisonnables d'usage** et de vente afin de s'assurer que le modèle économique proposé par *Europeana* ne repose pas sur des scénarios manifestement irréalistes.

Cette vérification chiffrée ne peut, à ce stade, qu'être effectuée sur des bases très approximatives compte tenu de l'incertitude importante liée aux modèles tarifaires qui seront retenus, et surtout à la demande potentielle. Elle vise simplement à montrer qu'**il est possible d'envisager sérieusement une rentabilisation des frais de numérisation** engagés par les éditeurs pour de nombreux ouvrages. Il ne s'agit pas de faire des budgets prévisionnels mais plutôt de proposer une « expérience de pensée » qui permette à ce stade de faciliter la discussion avec certains éditeurs hésitant à initier un processus de numérisation pour des raisons économiques.

On peut dans cet objectif proposer la simulation économique suivante :

- **Concernant les coûts techniques** associés au processus complet de vente d'une publication numérique, il faut distinguer plusieurs composantes :

- des **coûts de numérisation**, pour lesquels nous retiendrons les deux chiffres moyens avancés plus haut :
 - 15€par livre à numériser à partir de sources en papier
 - 60€par livre à convertir à partir de sources numériques
- des **coûts de mise en ligne** par le diffuseur du fichier et des métadonnées, qu'on peut estimer entre 5€et 15€par titre selon les protocoles d'échanges de

- données avec l'éditeur et les processus utilisés (qui ne sont pas entièrement automatisés) ;
- des **coûts de stockage** des fichiers (sachant par exemple qu'un fichier au format Adobe PDF protégé doit être stocké 3 fois sur les serveurs de DRM), différents selon qu'il s'agit de fichiers texte ou image (plus lourds). Par exemple, pour un fichier de 30Mo, on peut évaluer le coût à environ 1€par an ;
 - des **coûts de bande passante** utilisée pour la lecture en ligne ou pour le téléchargement des fichiers ; ces coûts peuvent apparaître très faibles, si on considère le poids des fichiers de livres numériques, comparativement par exemple aux fichiers vidéos, pour lesquels ils représentent une réelle contrainte du modèle économique. Néanmoins, ils peuvent devenir significatifs si des communautés importantes d'utilisateurs accèdent à des textes en ligne dans le cadre d'abonnements. Par exemple, le coût total d'un seul accès quotidien de 10 000 usagers à des textes en ligne peut être évalué à environ 1 000€par mois.
 - des **coûts d'utilisation des solutions DRM** ;
 - des **frais bancaires** sur les transactions, d'autant plus élevés que le montant des transactions est faible : de l'ordre de 3% pour les paiements par carte bancaire nationale et un montant de 10€ ils peuvent atteindre 20% à 30% pour les solutions de micro paiement, indiquées pour le règlement du « pay-per-view » ;
 - enfin, des **coûts d'assistance et de service après vente**.

Certains de ces coûts sont de purs coûts fixes initiaux, comme la numérisation et la mise en ligne, et parfois les solutions DRM, d'autres sont des frais annuels, comme le stockage, d'autres enfin des coûts variables, comme les coûts de bande passante, certains frais DRM, les frais bancaires et le SAV. A part la numérisation, en principe à la charge des éditeurs, la plupart des autres coûts sont en général assumés par les diffuseurs numériques et les librairies numériques au terme de contrats en contrepartie desquels ils gagnent une marge sur les ventes réalisées. A ces coûts purement directs et techniques s'ajoutent par ailleurs des **frais de promotion et d'acquisition de clientèle** à la charge de ces distributeurs.

Plutôt que de faire des hypothèses sur chacun de ces coûts, nous nous placerons **du point de vue des éditeurs** en supposant :

- qu'ils financent les frais de numérisation ;
- qu'ils reçoivent éventuellement des aides pour certains ouvrages ;
- qu'ils s'accordent avec les diffuseurs-distributeurs numériques sur un taux de marge, retenu par hypothèse en moyenne à 50%, couvrant l'ensemble des autres frais.

- **Concernant les produits de la commercialisation** des livres numériques, on peut faire l'hypothèse d'une solution combinant trois systèmes de tarification :

- la vente à l'unité de livres complets ;
- la vente de durées de consultation en « pay-per-view » ;
- la vente d'abonnements à des collections thématiques d'ouvrages.

La combinaison optimale de ces solutions tarifaires suppose de définir un **équilibre** entre les différents niveaux de tarifs afin d'**éviter les effets d'aubaine**, c'est-à-dire l'attractivité économique d'une des solutions manifestement plus avantageuse que les autres. Il devrait

également normalement importer aux éditeurs, d'une part d'avoir des prix concurrentiels et adaptés au marché, d'autre part d'éviter les dumpings dangereux pour l'équilibre à long terme du secteur. On peut en première approximation définir les contraintes suivantes, que devrait respecter cet équilibre, en nous faisant gré des simplifications destinées à la modélisation. Soient :

L : la durée de lecture moyenne d'un livre complet, notamment s'il est téléchargé

C : le temps de consultation cumulé moyen d'un livre numérique (en ajoutant la série des consultations en ligne courtes dont il fait l'objet).

L et C peuvent être différenciés si on considère qu'on n'utilise pas toujours un livre numérique, notamment en ligne, pour le lire du début à la fin, mais plutôt pour le consulter en y recherchant des informations.

T : le temps de lecture mensuel total moyen qu'un abonné consacre à la lecture des livres numériques.

N : le nombre moyen de livres qu'aurait acheté une personne ou une institution s'il n'y avait pas d'abonnement.

Q : le nombre de livres total d'une collection en abonnement.

d : la durée maximale autorisée de la consultation en pay-per-view.

P : le prix public TTC du livre numérique.

V : le prix public TTC de la consultation en pay-per-view.

A : le prix public TTC de l'abonnement mensuel.

Selon qu'on considère que le temps de lecture moyen est plutôt L ou C, on peut alors **définir quelques contraintes que devrait respecter une solution tarifaire équilibrée :**

(i) $V_1 > P.d / C$ ou $V_2 > P.d / L$

Cette contrainte assure qu'il n'est pas plus intéressant, si on souhaite lire ou consulter un livre complètement, de l'acheter par séries de consultations ponctuelles que de l'acheter d'emblée en totalité.

(ii) $A < P.T / C$ ou $A < P.T / L$

Cette contrainte assure qu'il existe une incitation à l'abonnement dans la mesure où il revient moins cher à un usager de s'abonner que d'acheter un par un les livres qu'il lira ou consultera.

(iii) $A > N.P$

Cette contrainte assure a contrario qu'il n'est pas systématiquement plus intéressant de s'abonner que d'acheter des livres à l'unité en respectant son budget habituel lorsqu'on est un acheteur de livre susceptible de s'abonner, ce qui diminuerait ipso facto la valeur totale du marché de ces livres. La valeur de N est évidemment difficile à fixer et dépend de la cible visée : individus, entreprises, laboratoires de recherche.

Q n'intervient pas dans ces contraintes, car il est évident qu'on ne peut pas fixer une valeur d'abonnement concurrentiel en totalisant d'une manière ou d'une autre les prix des livres constituant la collection. Le raisonnement consiste plutôt dans ce cas à penser que tous les livres de la collection n'auraient pas été achetés séparément, et que ce qui est offert est une possibilité de choix parmi des livres, associée à un nombre moyen estimé de consultations.

Q intervient par contre pour définir le revenu moyen attribué à chaque livre constituant la collection – en faisant une hypothèse de pro-ratisation linéaire.

Pour valider le réalisme économique potentiel de la vente de livres numériques, il faut alors **fixer des valeurs repères** pour chacune de ces variables. Le mini-modèle présenté doit encore une fois à ce stade être considéré comme un exercice approximatif, ou comme **un simple outil de simulation, chaque éditeur pouvant utiliser le modèle pour le « faire tourner » avec les valeurs réelles pertinentes pour ses livres et les hypothèses auxquelles il croit**, afin d'éclairer ses propres choix d'investissement dans la numérisation.

Voici des hypothèses qui pourraient être utilisées à titre d'exemple, en choisissant un ouvrage numérique dont le prix est de 10 €:

1. HYPOTHESES

nombre de livres d'un package abonnement	Q	500
nombre de livres moyen acheté par mois si pas d'abonnement	N	2 livres
temps moyen de lecture numérique par mois	T	30 heures
temps de lecture complète d'un livre	L	10 heures
temps de consultation moyen cumulé d'un ebook	C	2 heures
durée de pay per view autorisée	d	0,5 heures
prix TTC de l'ebook	P	10,00 €

Voici les intervalles de tarifs d'équilibre pour le pay-per-view et pour l'abonnement auxquelles ces hypothèses conduisent si on applique les contraintes (i), (ii) et (iii) :

2. EQUILIBRES TARIFAIRES

prix minimum 1 - pay per view	V1	2,50 €
prix minimum 2 - pay per view	V2	0,50 €
prix maximum 1 - abonnement mensuel	A1	30,00 €
prix maximum 2 - abonnement mensuel	A2	150,00 €
prix minimum abonnement mensuel	Amin	20,00 €

Considérons maintenant séparément le cas des livres, en général récents, qui feraient l'objet d'une adaptation ou d'une conversion numérique à partir de sources numériques. Le coût fixe moyen de numérisation à couvrir est de 60€ Supposons par ailleurs :

- que la TVA sur les livres numériques est de 19,6%, ce qui est le cas actuellement
- que la marge de diffusion-distribution est de 50% en moyenne
- que les droits d'auteur à régler par les éditeurs sont de 10% en moyenne

Si les livres numériques sont uniquement vendus à l'unité, **le point mort pour rentabiliser les frais de numérisation s'élève dans ce cas à 18 exemplaires par livre.**

Si les livres numériques sont vendus selon les trois modèles tarifaires proposés, il y a évidemment une infinité d'autres points d'équilibre. En supposant que le modèle pay-per-view, moins cher, et le modèle d'abonnement, attractif pour les chercheurs ou les professionnels, fonctionnent mieux que la vente traditionnelle à l'unité, on obtiendrait par

exemple un équilibre avec des valeurs telles que les suivantes, les tarifs se situant dans les intervalles définis par le tableau 2 :

TVA 19,6%

3. Point mort vente d'un livre si conversion numérique - exemple

	prix TTC	quantité	CA HT
pay per view	1,50 €	70	87,79 €
achat	10,00 €	5	41,81 €
abonnement annuel*	360,00 €	35	21,07 €
TOTAL CA HT			150,67 €
Frais de diffusion - distribution			50%
Droits d'auteur			10%
TOTAL MARGE HT			60,27 €
<i>*nombre de livres par collection en abonnement</i>	<i>500</i>		
Coût brut de numérisation			60,00 €

Considérons finalement le cas des livres, en général anciens, qui feraient l'objet d'une numérisation stricto sensu à partir de sources papier. Le coût fixe brut moyen de numérisation à couvrir est de 15€ Nous nous intéresserons au cas où les éditeurs bénéficient d'aides à la numérisation de ces livres avec une hypothèse d'un financement de 50%, soit un coût net de 7,50€/par livre. Les autres hypothèses restent inchangées.

Si les livres numériques sont uniquement vendus à l'unité, **le point mort pour rentabiliser les frais de numérisation s'élève cette fois à 3 exemplaires par livre.**

Si les livres numériques sont vendus selon les mêmes tarifs que dans l'exemple précédent, on obtiendrait par exemple un équilibre avec des valeurs telles que les suivantes :

TVA 19,6%

4. Point mort vente d'un livre si numérisation papier - exemple

	prix TTC	quantité	CA HT
pay per view	1,50 €	10	12,54 €
achat	10,00 €	0	- €
abonnement annuel*	360,00 €	10	6,02 €
TOTAL CA HT			18,56 €
Frais de diffusion - distribution			50%
Droits d'auteur			10%
TOTAL MARGE HT			7,42 €
<i>*nombre de livres par collection en abonnement</i>	<i>500</i>		
Coût net de numérisation			7,50 €

Les conclusions qu'on peut tirer de cette analyse sont les suivantes :

- Avec un modèle tarifaire unique, et encore plus à trois composantes, les performances de vente à réaliser pour rentabiliser les coûts de numérisation des livres ne semblent pas déraisonnables ou hors de portée. Et ce d'autant plus **que cette rentabilisation pourrait s'effectuer sur plusieurs années.** En ce qui concerne les livres récents, si les éditeurs choisissent les titres en fonction des indications du marché, vendre 18 exemplaires numériques ou alternativement et par exemple 5 exemplaires, 70 consultations et 35

abonnements à une collection de 500 titres incluant ce livre semblent représenter des ordres de grandeur tout à fait réalistes. En ce qui concerne les livres plus anciens dont les perspectives de vente sont plus faibles ou s'étalent sur un plus grand nombre d'années, les coûts nets de la numérisation aidée abaissent encore fortement ce point mort.

- Le modèle présenté et les valeurs envisagées ne tiennent pas compte d'autres facteurs susceptibles d'avoir un impact sur la rentabilité de la numérisation :

- L'adaptation des modalités tarifaires, notamment d'abonnement, aux usages collectifs : bibliothèques, entreprises.
- L'effet potentiel du référencement dans *Europeana* sur la vente de livres imprimés, qui pourrait être significative.
- Les coûts supplémentaires induits dans certains cas par le désarchivage ou par la nécessité d'acheter une version papier d'un livre indisponible sur le marché de l'occasion (qui pourrait toutefois être limitée par la collaboration de bibliothèques).

- Au-delà des valeurs unitaires de points morts, il faut considérer la somme en valeur absolue que peut représenter l'investissement dans la numérisation d'une collection de livres avant sa rentabilisation. De ce point de vue, seuls les éditeurs pourront décider des sommes annuelles qu'ils souhaitent consacrer à cette stratégie, en tenant compte de l'évolution du marché du numérique et de l'incitation à la vente de livres en général que représenterait *Europeana*. En ce qui concerne les livres pouvant bénéficier d'aides à la numérisation, il importerait par ailleurs de **chiffrer le montant d'investissement public que représenterait une aide à la numérisation de segments importants de livres sous droits** publiés dans les 70 dernières années, en fonction des thématiques envisagées par la politique documentaire. Ce chiffrage requiert une étude complémentaire.

5 Entretiens, acceptabilité et variantes

5.1 Entretiens réalisés

Conformément aux demandes du cahier des charges, l'étude a été menée en s'appuyant sur un certain nombre d'entretiens avec des acteurs représentatifs du secteur du livre dans son ensemble, et certains spécialistes. Un nombre important d'entretiens a été en particulier effectué avec des acteurs français, dont l'acceptation est le principal facteur clé de succès pour le démarrage du projet *Europeana*. Voici une liste des personnalités qui ont été directement consultées au cours d'entretiens en face à face :

→ Editeurs :

- Piero ATANASIO, Membre de la direction de l'Association des Editeurs Italiens
- Anne BERGMAN-TAHON, Directrice de la Fédération des Editeurs Européens
- Jean-Manuel BOURGOIS, PDG des éditions Magnard-Vuibert
- Alban CERISIER, Responsable du pôle numérique des éditions Gallimard

- Patrick FENOUIL, PDG des éditions Hermès Science-Lavoisier et de la librairie Lavoisier
- François GEZE, PDG des éditions La Découverte
- Sylvie MARCE, Directrice du développement des éditions Belin
- Christine de MAZIERES, Secrétaire Général du Syndicat National de l'Édition
- Arnaud NOURY, PDG du groupe Hachette
- Michel PRIGENT, Président des Presses Universitaires de France

→ **Bibliothécaires :**

- Catherine FORESTIER, Responsable du pôle « livres électroniques » du consortium Couperin (Établissements d'Enseignement Supérieur)
- Alain PATEZ, Responsable de la Bibliothèque Numérique pour le Handicap et de la Bibliothèque Numérique de la Ville de Boulogne-Billancourt
- Vigdis Moe SKARSTEIN, Directrice de la Bibliothèque Nationale de Norvège

→ **Libraires :**

- Hélène CLEMENTE, Responsable du projet de portail numérique du Syndicat Français de la Librairie
- Françoise DUBRUILLE, Directrice de l'European Booksellers Federation
- Eric HARDIN, Président de ALIRE (Association des Librairies utilisatrices de Réseaux Electroniques)
- Thierry LECOMTE, Directeur Commercial de l'Appel du Livre, Administrateur de ALIRE
- François MILLIET, Libraire, administrateur de ALIRE
- Sophie de SAINT MARC, Juriste du Syndicat Français de la Librairie

→ **Auteurs :**

- Alain ABSIRE, auteur, Président de la Société des Gens de Lettres
- Myriam DIOCARETZ, auteur, Secrétaire Général du European Writers Congress

→ **Autres experts :**

- Bernard de FREMINVILLE, Directeur Général de Dilicom
- Alain PIERROT, Business Development Manager chez i2S, spécialiste d'édition numérique, co-auteur du livre *Le monde du livre en question* (Actes Sud, 2001)
- Bernard WALLISER, économiste, Professeur à l'École des Hautes Études en Sciences Sociales et à l'École Nationale des Ponts et Chaussées.

A ces personnalités il faut ajouter tous les membres du groupe de travail BnF/SNE devant lesquels une première présentation a eu lieu et plusieurs responsables du projet *Europeana* au sein de la BnF, sous la responsabilité de Agnès SAAL, Directrice Générale :

- Valérie GAMME, Directrice du Service Juridique

- Denis BRUCKMAN, Adjoint au Directeur des Collections pour les questions scientifiques et techniques
- Stéphane PILORGET, Responsable du Bureau Conception et Développement, Direction des Services et des Réseaux, Département des Systèmes d'Information, Service Etudes et Développement
- Lucien SCOTTI, Directeur des Affaires européennes et internationales, Délégation aux relations internationales.

Enfin, deux présentations publiques du modèle économique recommandé ont été effectuées devant les membres des bureaux exécutifs d'organisations représentant les éditeurs et les libraires :

- La Fédération des Editeurs Européens
- La European Booksellers Federation

5.2 Acceptabilité

L'essentiel des remarques, critiques ou suggestions issues de ces entretiens et qui nous ont semblé pertinentes sont incluses dans les recommandations faites dans le présent document. On peut souligner en complément les points saillants suivants :

- Une très large acceptabilité des **éditeurs** sur la base du modèle économique recommandé, qui permettrait de collaborer à *Europeana* dans des conditions juridiques claires avec des conditions de rémunération acceptables et des formes d'organisation commerciales souples. Les éditeurs rencontrés souhaitent que soit reconnue la valeur de l'apport à *Europeana* des éléments permettant à celle-ci de présenter les métadonnées de leurs ouvrages sur son site et d'effectuer une recherche plein texte au sein de ces ouvrages. Au-delà du modèle économique général, ils souhaitent que cet apport ait des contreparties, dont deux sont essentielles :

- un **signalement spécifique des ouvrages** sous droits au sein du site Europeana, associée à une mise en valeur à travers les métadonnées riches appropriées ;
- tant que faire se peut, des **aides publiques à la numérisation** des ouvrages plus anciens susceptibles d'intéresser la Bibliothèque numérique européenne mais dont les perspectives de rentabilisation commerciale à court terme sous forme numérique sont plus incertaines.

Enfin, les éditeurs souhaitent en général conserver l'intégrale maîtrise des fichiers numérisés, de leur hébergement dans toute la mesure du possible (d'où l'intérêt de protocoles tels que Open Search™), et de leurs modes de commercialisation.

- Une bonne acceptabilité des **libraires** dans la mesure où leur inclusion dans le modèle économique général est clairement actée et mise en œuvre, sur la base d'une définition de la notion de « librairie », et avec une vigilance et un intérêt particulier apportés aux

modalités techniques de liens vers leurs sites et vers les fiches produits des ouvrages, en particulier imprimés.

- Une volonté de la part des **associations d'auteurs** d'être considérées comme des parties prenantes d'un projet d'une telle ampleur, dont elles acceptent le principe, en étant en particulier directement informées de ses avancées et modalités. Et une vigilance apportée à la nécessité de s'assurer, dans toutes les conditions légales et contractuelles appropriées, de l'accord des ayants droit des ouvrages présentés et diffusés en version numérique, avec une volonté de collaborer à la mise en place d'outils facilitant ce processus.

- Les **bibliothèques** autres que nationales ne sont pas directement incluses dans le modèle économique proposé, pour des raisons évoquées, hormis celles spécialisées dans le handicap. Cela ne semble pas être un inconvénient à ce stade, dans la mesure précisément où *Europeana* ne jouerait pas, pour les livres sous droits, un rôle qui aurait pu être concurrentiel de ces bibliothèques ; mais au contraire pourrait représenter un outil de travail intéressant pour elles et une ressource documentaire en ligne importante à signaler à leurs lecteurs.

5.3 Variantes

Toutefois, il ressort des entretiens, en particulier en dehors de France, qu'il serait intéressant de considérer dans un deuxième temps des **variantes** au modèle central recommandé par l'étude. En particulier :

- Dans d'autres pays européens, notamment les pays scandinaves, les bibliothèques nationales pourraient avoir une préférence pour la négociation d'accords collectifs avec les éditeurs et les ayants droits afin d'acquérir des droits de présentation gratuite au public de fonds d'ouvrages numérisés, sous la forme d'accès illimités en ligne. Autrement dit, elles pourraient avoir une **préférence pour les modèles économiques de « bibliothèque numérique » de type B6 ou B8**. C'est sur cette base que la Bibliothèque Nationale de Norvège a récemment négocié un accord pour la consultation libre en ligne d'un millier de livres ayant trait à l'histoire et à la culture norvégienne. En contrepartie d'une rémunération fixe globale et d'une prise en charge intégrale de la numérisation. Un tel modèle économique n'est pas incompatible avec le modèle principal que nous recommandons : il pourrait le compléter quand des bibliothèques nationales concluent de tels accords. En effet, dans ce cas, les livres sous droits ainsi négociés pourraient simplement être traités comme des livres du domaine public. Alternativement, si l'accord implique que la consultation se fasse dans certaines conditions de protection sur le site de la bibliothèque nationale de référence, celles-ci joueraient pour ces livres le rôle joué par ailleurs par les diffuseurs agréés. Remarquons toutefois que l'accord signé en Norvège exclut tout téléchargement, et porte sur des livres peu exposés commercialement : il ne semble pas qu'il soit facilement extensible à d'autres formes de lecture numérique et à d'autres types d'ouvrages en grande quantité.

- La variante précédente pose d'une manière plus générale la question du **circuit des relations entre *Europeana*, les bibliothèques nationales et les détenteurs de droits**. Dans le modèle présenté, nous avons considéré seulement deux catégories d'acteurs : *Europeana* d'une part, les diffuseurs agréés représentant les éditeurs d'autre part. Devrait-on considérer *Europeana* comme une entité négociant directement avec les diffuseurs de chaque pays, ce qui permettrait d'envisager une croissance rapide de son fonds en collaboration avec les acteurs privés intéressés et sans médiatisation par des décisions publiques nationales ? Ou devrait-on plutôt envisager *Europeana* comme une fédération de fonds de livres apportés par les différentes bibliothèques nationales, négociant chacune avec les éditeurs de leur pays, au fur et à mesure que des décisions publiques seront prises en ce sens ?... Nous nous contenterons de soulever cette question, dont la réponse n'est pas seulement de nature économique.

- Certains éditeurs d'autres pays pourraient considérer a contrario que le modèle recommandé est trop structuré et contraint, de par le fait qu'il suppose une gestion des accès aux livres numériques par des diffuseurs agréés uniques. Si *Europeana* doit être un portail de recherche pour les livres sous droits, il faudrait envisager plutôt que les métadonnées présentées permettent à l'utilisateur de se diriger librement vers une liste exhaustive de sites commercialisant sous des formes diverses l'ouvrage trouvé, par exemple en exploitant certaines possibilités futures de la norme DOI. Cette variante mérite d'être considérée et discutée avec les éditeurs, mais elle nous semble soulever pour le démarrage un certain nombre de difficultés techniques et de principe : les DOI ne sont pas d'un usage courant et les possibilités considérées sont encore à l'état de projet, la responsabilité de la fourniture à *Europeana* des données n'est plus claire dans ce schéma, la navigation des internautes risque d'être rendue plus complexe et le respect de principes déontologiques communs plus difficile à établir.

6 Mise en œuvre et compléments d'étude recommandés

L'étude telle qu'elle est présentée dans ce document répond aux exigences du cahier des charges élaboré par la BnF en vue de la proposition d'un modèle économique de participation des éditeurs à la bibliothèque numérique européenne. Les recommandations qu'elle contient s'appuient sur des considérations générales et théoriques préalables et formulent des principes opérationnels pour définir ce modèle économique. Au-delà, l'étude propose des hypothèses concernant les schémas techniques d'implémentation de ce modèle économique. Les recommandations formulées sont de nature à envisager dans de brefs délais une phase de validation et de négociation avec des éditeurs français acceptant ce modèle et désireux de collaborer au projet *Europeana*.

Plusieurs points de l'étude conduisent toutefois à suggérer de mener à bien des analyses complémentaires afin que cette collaboration se déroule dans les conditions les plus claires possibles. Ces points concernent en particulier :

- la validation et la précision des normes de métadonnées et de fichiers d'indexation, ainsi que des protocoles de transferts de ces métadonnées et d'accès à ces fichiers ;

- la validation de la nécessité d'une politique documentaire associée à la possibilité pour les éditeurs de bénéficier d'aides publiques à la numérisation de certains fonds éditoriaux, et, si cette validation a lieu, la définition précise de cette politique ;
- le chiffrage des coûts de numérisation impliqués par la politique documentaire pour les ouvrages à numériser à partir de sources papier, ou d'une manière plus générale pour tous les ouvrages sous droits dans un certain nombre de segments thématiques.

Par ailleurs, les capacités d'entraînement du projet *Europeana* dans le reste de l'Europe seraient accrues par une analyse plus approfondie de l'acceptabilité du modèle en dehors de la France, tant auprès des bibliothèques nationales que des éditeurs, et la considération de l'ouverture potentielle du modèle à des variantes permettant de tenir compte à la fois des spécificités nationales, et des évolutions des technologies et des normes au service des publications numériques.

* *

*