

QUELLE FILIÈRE INDUSTRIELLE POUR LA NUMÉRISATION DU PATRIMOINE ?

LE 1 JUILLET 2010 CHRISTIAN FAURÉ

Quelle voie intermédiaire prendre, entre une logique libérale de la privatisation basée sur la publicité et visant uniquement la rentabilité, et une politique culturelle sous perfusion publique, mais à perte ?



Le contexte de la consultation sur le Grand Emprunt

La cacophonie et la mécompréhension autour du Grand Emprunt, et plus précisément sur le volet numérisation, font qu'on est actuellement dans une situation de crise, au sens propre du terme : **quelque chose va se décider**.

Au départ, c'étaient 150 millions qui devaient être alloués aux institutions pour qu'elles puissent poursuivre et accélérer les projets de numérisation ; au final ce ne sont plus que des montants de prêts (donc remboursables avec intérêts) pour favoriser la mise en place d'une filière industrielle du numérique, basée sur des partenariats publics/privés.

On sait que l'actualité de la crise économique de ces derniers mois a certainement beaucoup favorisé la **formulation très libérale** de la consultation publique (le développement du « **machin numérique** ») lancée par le secrétariat de la Prospective et du Développement de l'économie numérique. De plus, dans le cadre d'une période d'austérité et de restrictions budgétaires importantes dans les dépenses de l'État, le Grand Emprunt devient un dossier beaucoup particulièrement épineux pour le gouvernement : difficile de dire « on fait les valises et on rentre » après avoir fait de la relance par l'innovation un axe important de la stratégie française.

Deux tentations s'opposent donc entre celle du ministère de la Culture et celle du ministère des Finances : le premier veut continuer à croire à la nécessité d'une politique culturelle tandis que le second tente de radicaliser les choix qui devront être faits sur la base exclusive du principe de rentabilité. Il n'y a donc **plus de consensus au sein même du gouvernement sur l'avenir du Grand Emprunt**, et les différentes institutions qui doivent participer à la solution (BnF, bibliothèques municipales, INA, IRCAM, Cinémathèque, Cité des Sciences, archives, musées, etc.) ne comprennent plus la règle du jeu, qui semble par ailleurs changer chaque jour en ce moment.

La vision qui est présentée ici est une **tentative de réponse à la consultation publique sur le volet numérique**. Elle a l'ambition de sortir par le haut des apories dans lesquelles la question de la numérisation du patrimoine dans le cadre du grand emprunt se retrouve aujourd'hui.

La publicité est-elle la solution ?

L'activité industrielle autour de la numérisation de contenus culturels et patrimoniaux est **l'activité de numérisation qui est aujourd'hui la moins rentable** si on la compare aux archives, cadastres et autres documents administratifs (littérature grise). D'autre part, on sait que Google a beaucoup investi sur cette activité avec sa plate-forme **Google Books dont on commence à peine à entrevoir l'ampleur**. Quel industriel voudrait, dans ces conditions, prendre le risque d'investir sur un secteur d'activité à faible potentiel rémunérateur tout en ayant la machine de guerre de Google en embuscade ? Soyons clairs : personne. Il faut donc poser le problème différemment.

Commençons pour cela par **évacuer toutes les fausses bonnes idées** que l'on peut entendre sur le modèle d'affaire qui pourrait rendre cette filière numérique rentable. Pour cela il faut d'abord savoir que la numérisation d'un ouvrage n'est, en moyenne, **rentabilisée qu'au bout de vingt ans**, uniquement en ce basant sur le service de reproduction que propose la BnF. C'est une moyenne car, bien évidemment, certains ouvrages ne font l'objet d'aucune demande de reproduction. Quand se pose la question de savoir comment ce seuil peut être abaissé ne serait-ce que sur dix années, la réponse que j'entends systématiquement est : **la publicité**.

La publicité est généralement le joker que l'on avance quand on est à court d'idées. Et c'est assurément le modèle d'affaire le plus simple à proposer : il me manque 100 millions ? Qu'à cela ne tienne, la pub fera le reste. Comment et sur quelles bases ? La réponse est généralement plus évasive. Faut-il monter un mécanisme et une régie publicitaire en propre ? Faut-il s'appuyer sur les solutions clés en mains proposées par Google ? Cette dernière réponse serait pour le moins ironique puisque Google aurait une part importante du bénéfice publicitaire sans avoir investi dans la numérisation. Faire sans Google, c'est à l'inverse prendre le risque de se retrouver **dans le collimateur d'un industriel du web qui s'y connaît** et qui a les moyens de ses ambitions.

On préférera donc essayer de composer avec Google plutôt que de le concurrencer sur son propre terrain en faisant « Cocorico ! ». **Les arguments basés sur la valorisation via un modèle d'affaire fondé sur la publicité ne tiennent pas la route**, encore moins quand l'on sait que la valeur publicitaire sur le web, comme l'avait écrit Tim O'Reilly dès 2007, tend à se diluer très fortement. C'est la raison pour laquelle Google doit indexer toujours plus de contenus, nativement numériques ou à numériser, pour **amortir la baisse tendancielle de la valeur unitaire et nominale de la publicité**.

Que vaut le numérique ?

Retour à la case départ : comment valoriser la numérisation du patrimoine ? Songeons-y un instant, si l'on se donne tant de mal pour imaginer un modèle d'affaire viable pour une filière industrielle de numérisation, c'est peut-être parce que **le numérique, de manière tendancielle, ne vaut rien**. Le numérique a un coût, surtout lorsqu'on doit numériser, mais, une fois l'investissement réalisé, financièrement et en tant que tel, il ne vaut plus rien. **Soyons plus précis : un fichier numérique ne vaut rien**. Et c'est bien la raison pour laquelle le monde de l'édition freine des quatre fers lorsqu'il s'agit de faire circuler un fichier numérique existant (même pour en donner une copie pour archive à une institution, la plupart refusent). **Un fichier numérique en circulation, c'est de la nitroglycérine pour celui qui en attend une source de revenu**.

Acceptons donc cette thèse, qui est aussi une hypothèse de travail, que le fichier numérique ne vaut rien. Et vérifions cette proposition :

pour les institutions, c'est généralement le service de reproduction qui est la principale source de revenu, c'est-à-dire le retour à l'impression papier.

pour les plates-formes de diffusion de contenus numériques, on sait bien que ce n'est pas le fichier numérique que l'on paye mais un écosystème technologique (format de fichiers propriétaires, logiciels verrouillés, périphériques spécifiques, fonctionnalités d'achat rapide brevetées, etc.)

pour d'autres initiatives plus confidentielles mais notables (par exemple **PublieNet**), c'est la qualité d'une présence sur le web et la sensibilité de la communauté des lecteurs/clients qui fait la différence : entre l'éditeur numérique et les lecteurs/acheteurs, il y a un crédit et une confiance.

La valeur d'un fichier numérique a donc besoin d'un service autre que la simple diffusion

pour pouvoir avoir une valeur financière.

Le service de reproduction doit devenir le premier industriel d'impression à la demande

Loin d'enterrer les poussiéreux services de reproduction, il faut les muscler. Ces services, qui aujourd'hui nous semblent d'un autre âge, doivent se doter d'un **service d'impression à la demande digne des autres acteurs leaders sur ce créneau**. L'économie d'échelle qu'ils peuvent avoir, qui plus est sur la base d'oeuvres particulièrement attrayantes ne peut qu'être profitable. Cette re-fondation peut ramener dix ans, au lieu des vingt actuels, le délai d'amortissement d'une numérisation.

La chose n'est pas gagnée d'avance pour autant : il faut une **plate-forme web en self-service** qui demande du travail, il faudra être très rapide et **avoir une logistique aussi affûtée que celle d'Amazon**, a minima sur le territoire français. L'objectif est clairement de livrer au domicile d'un client l'impression d'un ouvrage relié de qualité en moins de 48 heures, et à peine plus s'il y a une demande d'impression personnalisée.

Sur cette voie, il va y avoir des **frictions avec les plate-formes de distribution des éditeurs de la chaîne du livre**. Mais pas dans l'immédiat puisque les modèles sont actuellement différents (pas d'impression à la demande, pas de self-service et pas de livraison au particulier), mais si la plate-forme d'impression à la demande est un succès, elle pourra proposer ses services différenciants aux éditeurs (traditionnels, mais aussi numériques) : par exemple proposer des « templates » de formats variés et personnalisables. N'oublions pas que près des trois quarts du coût d'un livre représentent les coûts d'impression, de distribution, de diffusion et de points de vente.

Le cas Gallica



La filière de numérisation peut donc trouver un premier modèle économique dans l'impression. Pour où l'on voit que la valorisation de la numérisation se fait d'abord sur... l'impression. Mais se pose toujours la question de **la diffusion sous format numérique et en ligne**. Premier constat : c'est la vocation de Gallica. On comprendra dès lors que la filière numérique qui est appelée de ses vœux par le gouvernement aura du mal à accepter de faire le travail de numérisation pour que le fruit de son investissement se retrouve diffusé en ligne gratuitement sur Gallica.

Gallica devra être repensée, et pour commencer **il faut que la bibliothèque numérique quitte le giron exclusif de la BnF**. Cela veut dire que Gallica aura le statut d'un établissement public-privé dans lequel l'ensemble de plate-forme technologique sera possédée et gérée par le consortium privé investissant dans la filière numérique.

Statutairement, **la BnF doit garder le contrôle et la maîtrise de la politique culturelle que**

porte Gallica. Mais cette maîtrise ne sera plus exclusive, elle devra être partagée car si cette bibliothèque en ligne se nourrit des ouvrages numérisés, et il faudra bien un *modus vivendi* et des droits de quotas pour chacun : la BnF peut vouloir numériser en premier des ouvrages qui ne sont pas jugés commercialement opportuns pour le partenaire privé. Un système de quotas, qui devra évoluer dans le temps, doit être mis en place. Par exemple, sur les cinq premières années, sur dix ouvrages numérisés, le partenaire privé pourra en choisir cinq, tout comme la BnF. Par la suite, **les résultats de la filière numérique serviront de référent pour faire évoluer les quotas** : si la filière est sur le chemin de la rentabilité le ratio peut s'infléchir en faveur de la BnF, ou l'inverse si la rentabilité tarde à se faire jour. L'essentiel est de ne pas figer la formule et d'y introduire une variable dépendant de la rentabilité, sans quoi tout l'édifice s'effondre.

Cette réorganisation du statut juridique de Gallica devra nécessairement initier une refonte de la politique de gestion des droits des oeuvres qui n'est pas opérationnelle en l'état actuel (une licence sur mesure que ne peuvent pas exploiter les robots, et que d'ailleurs personne ne comprend vraiment).

Bien évidemment, d'un point de vue technologique, **la plate-forme de service d'impression évoquée précédemment sera nativement intégrée à Gallica**, on peut même forcer le trait en disant que Gallica ne sera qu'un module de la plate-forme d'impression.

Les métadonnées : clés de voûte de la nouvelle filière industrielle

Aussi étonnant que cela puisse paraître, dans cette consultation publique sur « le développement de l'économie numérique », il n'y est **jamais question de métadonnées**. Le mot n'y apparaît même pas une seule fois le long des trente-neuf pages du document. C'est proprement sidérant. Et ça l'est d'autant plus que **la politique industrielle qui va être mise en place devra placer la question des métadonnées au cœur de tout le dispositif industriel**.

Si l'impression à la demande était **le volet diffusion papier** et Gallica **le volet diffusion numérique**, ces deux activités passent à un niveau supérieur grâce à la politique sur les métadonnées. **La richesse numérique de notre patrimoine est directement proportionnelle aux métadonnées qui le décrivent**. Le trésor des institutions patrimoniales réside aussi et surtout dans leurs catalogues et leurs thésauri : tout comme on ne peut gérer un patrimoine physique sans métadonnées la question devient encore plus urgente quand l'oeuvre est numérisée : **une politique numérique sans politique des métadonnées n'est qu'une chimère, un délire, une schwarmerei comme disait Kant**.

Plutôt que de me répéter, je vous renvoie ici à ma note sur **Les enjeux d'une bibliothèque sur le web** où il était question des **orages sémantiques** mais aussi d'étendre la pratique de gestion d'un catalogue d'oeuvres à **une pratique de gestion d'un catalogue des discussions et des polémiques relatives à ces oeuvres**. Ainsi, fort de ce nouveau positionnement, et sur la base de sa nouvelle plate-forme technologique, la nouvelle filière industrielle du numérique pourra proposer des outils avancés à l'Éducation nationale pour doter l'enseignement d'un outil d'annotation et de contribution qui dépasse la vision simpliste et fade des « like », et donne enfin le pouvoir aux enseignants d'enseigner.

Chaque plate-forme de diffusion des oeuvres numériques rencontre très vite sa limite **dans les faiblesses de sa politique des métadonnées**. Le cas d'iTunes est représentatif : c'est une panique monstre pour faire des découvertes dans le catalogue, c'est pourtant paradoxal quand on sait que, même sur iTunes, les métadonnées (titre, auteur, artistes, jaquette, etc.) sont **la vraie valeur des fichiers numériques** (Cf. **Quand les métadonnées ont plus de valeur que les données**).

Pour les oeuvres qui sont du ressort de la BnF, le travail de bascule de l'ensemble des catalogues au format du web sémantique avec leur diffusion sur le web a déjà été initié : **cette démarche est la clé de voûte, à la fois technologique et économique, de tout le système**. Pour les oeuvres audios et vidéos (des oeuvres de flux), les outils d'annotation contributives (avec des métadonnées BottomUp et TopDown) doivent **être développés en complément des catalogues descriptifs existants**.

Le catalogage des orages sémantiques permet également d'obtenir tout un appareil critique issu des informations collectées via le dispositif des orages sémantiques. Si celui-ci est géré par la BnF, on peut réussir à mener une politique industrielle des technologies numériques dont le coeur du dispositif s'appuie, et trouve son crédit, dans la politique culturelle. **Une logique économique exclusivement consumériste n'est pas une fatalité**, loin s'en faut, car ce qui est brièvement décrit ici est un chemin vers **une économie de la contribution** financièrement rentable.

On peut donc sortir de l'alternance destructrice entre :

d'un côté une logique libérable de la privatisation adossée à une vision exclusive sur les retours sur investissement à court terme, grâce au dieu de la publicité ;
de l'autre une politique culturelle maintenue sous perfusion publique, mais à perte (la logique de la réserve d'indiens).

Que le Grand Emprunt accouche de quelque chose ou non, nous n'échapperons pas à cette lancinante question : quelle politique industrielle pour les technologies de l'esprit ? La seule réponse crédible passe par le positionnement de la politique culturelle au cœur de l'outil industriel, pas à côté. « **Trade follows film** » disait le sénateur américain McBride en 1912 : on va peut-être arriver à le comprendre cent ans plus tard en France, notamment pour donner au commerce et à l'économie un autre visage que le consumérisme américain.

Enfin, par pitié, arrêtons de parler systématiquement de e-tourisme dès qu'il est question des territoires. **Les territoires sont autre chose que des destinations touristiques**, et les régions n'hivernent pas toute l'année pour se réveiller quand les Parisiens et les étrangers prennent leur vacances. Ces modèles d'affaire sur le e-tourisme sont dangereux et méprisants.

—

Billet initialement publié sur **le blog de Christian Fauré**

Images CC Flickr **Troy Holden** et **►bEbO**

CASSANDRE

le 6 juillet 2010 - 4:21 • SIGNALER UN ABUS - PERMALINK



Il est dommage que l'auteur soit d'une ignorance crasse sur la question des archives (qui ne sont pas assimilables à de la littérature grise) et que ce débat soit totalement confisqué par la numérisation du livre imprimé.

Comme d'habitude, les acteurs de l'imprimé sont si peu occupés dans leur domaine professionnel, qu'ils ont tout loisir d'agiter de petites idées creuses et de brasser du vent. Ils s'autoproclament de plus autorisés à parler au nom de l'universel et s'assoient ainsi avec ignorance, sinon avec mépris, sur d'autres enjeux, plus graves, plus fondamentaux que les "services de reproduction papier", et que renferment la question de la numérisation des archives.

C'est poser, à force d'ignorance et de mépris pour les archivistes dont personne ne souhaite écouter la voix et de ce dont ils sont dépositaires, le premier totalitarisme numérique, en tous temps, en tous lieux, dirigé contre tous les individus dont le nom et les informations personnelles apparaîtront dans les archives publiques numérisées.

Les petits débats des gens de l'imprimé apparaissent comme vraiment puérils face à un tel enjeu.

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

3 pings

Les tweets qui mentionnent *Quelle filière industrielle pour la numérisation du patrimoine ?* » Article » owni.fr, digital journalism -- Topsy.com le 1 juillet 2010 - 19:26

[...] Ce billet était mentionné sur Twitter par damien douani. damien douani a dit: Quelle filière industrielle pour la numérisation du patrimoine ? http://bit.ly/90MIOu #owni [...]

Patrimoine-en-blog » Blog Archive » Revue de presse juillet 2010 le 5 juillet 2010 - 5:57

[...] Quelle voie intermédiaire prendre, entre une logique libérable de la privatisation basée sur la publicité et visant uniquement la rentabilité, et une politique culturelle sous perfusion publique, mais à perte ? Lire la suite [...]

La part d'ombre de Google Livres « Le fil de la médiathèque le 16 juin 2012 - 21:31

[...][...]