

BNF ET GOOGLE : CONTRIBUTION AU DÉBAT

LE 12 JANVIER 2010 OLIVIER ERTZSCHEID

Vendredi dernier, se tenait à la BnF une journée de débat autour du futur-très-probable accord entre Google et ladite BnF, débat intitulé "Numérisation du patrimoine et moteurs de recherche" ...

Vendredi dernier, se tenait à la BnF une journée de débat autour du futur-très-probable accord entre Google et ladite BnF, débat intitulé "**Numérisation du patrimoine et moteurs de recherche**", débat réunissant notamment Patrick Bazin, Roger Chartier et Robert Darnton, soit quelques analystes et acteurs les plus passionnants sur le sujet. Le programme complet de cette journée était initialement disponible sur le site de la BnF (**ici**) mais, étrangement, il a été "soit déplacé soit supprimé". :-)

En ligne, très peu de traces de cette journée : merci à l'**ADBS** et à **Livres Hebdo** d'avoir assuré de rapides compte-rendu (si vous en avez repéré d'autres, les commentaires sont ouverts). Très peu de traces ... et pour cause : les personnes présentes sur place n'ont même pas pu twitter étant donné l'absence de connexion Wifi ou 3G.

Plus globalement, et alors même que depuis les plus hautes sphères de l'État jusqu'à la presse grand public, chacun se passionne pour ce sujet (ce qui n'est pas toujours le cas lorsque l'on parle de l'avenir et du modèle des bibliothèques), plus globalement donc, le fait que les organisateurs de cette journée n'aient pas envisagé de la webcaster ou d'en diffuser des extraits est atterrant.

Et à peu près aussi improbable que TF1 décidant d'organiser la soirée de l'élection de Miss France à huis-clos. C'est vrai quoi, on a un sujet passionnant, les meilleurs spécialistes (pro ou anti) de la question sur le plateau, une demande très forte d'information de la part du public, et on la joue "black-out", on l'organise sur le mode "et ben vous aviez qu'à être présents".

Domage. Dommage. Dommage (et après on s'étonnera et on n'en finira pas de gloser sur l'incapacité des bibliothèques françaises à communiquer ... **question de culture ?**). Bon, on va espérer que dans un sursaut de lucidité la BnF ait finalement filmé l'intégralité de la journée et qu'elle se prépare à la mettre en ligne sur son site :-)

"Vous avez été nommé voilà deux ans de cela à la tête de la Bibliothèque Nationale de France (BNF)"

De mon côté, vendredi dernier, j'étais en train de corriger un paquet de copies. Petite contribution au débat donc, mais, pour une fois sur ce blog, cette contribution ne sera pas la mienne :-). Plutôt celle de mes étudiant(e)s en deuxième année **métiers du livre à l'IUT de La Roche sur Yon**. Après une petite vingtaine d'**heures de cours passées sur le sujet de la numérisation**, ils et elles ont eu les vacances de Noël pour plancher sur le sujet suivant :

"Vous avez été nommé voilà deux ans de cela à la tête de la Bibliothèque Nationale de France (BNF). Voilà un an que vous travaillez sur le dossier de la numérisation patrimoniale et sur la possibilité d'un accord avec Google. Le 7 Janvier 2010, vous prononcez un discours à l'assemblée nationale pour faire part de votre décision sur ce dossier.

En vous inspirant de l'actualité sur le sujet et en vous appuyant sur les éléments vus en cours vous rédigerez un discours indiquant si oui ou non vous avez décidé de travailler avec Google en mentionnant – en cas de réponse négative – quels sont vos arguments, et – en cas de réponse positive – quelles sont vos conditions.

Le devoir devra être impérativement manuscrit et ne devra pas faire moins qu'une copie double (4 feuillets)."

Résultat des courses : sur 26 étudiant(e)s, 16 copies pour un accord avec Google, et 10 contre. Naturellement, l'honnêteté m'oblige à indiquer que si leur professeur avait été différent, la proportion de leurs réponses aurait sensiblement pu varier :-)

Du côté des "contre", les arguments les plus massivement retenus sont les suivants :

problème de l'opacité de l'algorithmie de la recherche de livre (= pourquoi tel livre et non tel autre en première position ? en fonction de quel critère ? Tant que l'on se saura pas comment ça marche, on ne signera pas)

problème de la représentativité (par rapport à la **charte des bibliothèques, article 7**)

problème de l'exclusivité (sur un temps variable mais globalement assez long) et de la

“stérilisation” de la copie numérique remise à la bibliothèque (pour le même temps assez long)

problème de l'exclusivité d'indexation (métadonnées non-accessibles aux autres bib/moteurs)

problème de la compétence (= la numérisation doit être, rester et re-devenir) une compétence-coeur du monde des bibliothèques. Accepter de déléguer, c'est prendre le risque de perdre cette compétence.

problème du monopole : “risque que Google devienne LE catalogue de référence.”

Du côté des “pour”, voici les raisons et les exigences retenues comme “non-négociables” pour finaliser l'accord :

Les raisons tout d'abord :

L'urgence et l'argent : “*Nous devons accepter de déléguer à Google l'équivalent de nos numérisations annuelles. Cela rendra nos dépenses quasi-nulles (...) ce qui nous permettra d'élaborer un projet d'envergure dans les décennies à venir. (...) afin de développer en parallèle des structures de numérisation d'envergure au niveau national et européen.*” Cette importance de **structures (d'infrastructures en réalité) de numérisation alternatives** me semble être le coeur du sujet. Dans une autre copie, Gallica est citée comme pouvant servir (à condition d'être financée et développée en conséquence) d'infrastructure type.

Les exigences ensuite (venant de copies différentes, certaines des exigences listées ci-dessous sont contradictoires) :

Opt-in. Étrangement (il me semblait avoir montré la difficulté voire l'inanité de la chose ...) les étudiant(e)s sont “dans la loi” tout autant que dans la crainte d'une entorse au droit d'auteur tel qu'il existe aujourd'hui en France. Presque tous ceux qui acceptent de signer avec Google réclament donc qu'il respecte, pour les oeuvres orphelines, une procédure d'Opt-In en lieu et place de l'Opt-Out actuellement en vigueur. “*Nous imposons à Google de ne diffuser les ouvrages numérisés qu'après avoir obtenu l'autorisation des ayant-droit.*” Il y a peu de chances que cela soit le cas, mais il est intéressant de noter que même après mon bourrage de crâne volontariste, cette “option” reste plébiscitée :-)

Money. “*La BnF a exigé et obtenu que 60% des bénéfices tirés de ces ventes (sur les ouvrages numérisés depuis la BnF) lui revienne.*”

Métadonnées : “*Il serait nécessaire que chaque bibliothèque partenaire, lorsqu'elle prête l'ouvrage en vue de sa numérisation, fournisse les méta-données. les exemplaires numérisés seraient alors indexés d' manière fiable.*” Noble proposition (déjà en partie réalisée, Google allant s'abreuver en métadonnées dans différents catalogues “ouverts”) mais qui risquerait, in fine, de fournir à Google le beurre, l'argent du beurre et une ristourne sur l'addition comme le dit le proverbe.

No-logo. “*Les caractéristiques des ouvrages numérisés doivent être les mêmes pour tous : pages en noir et blanc, couverture en couleur de 200 à 400 dpi, format jpg ou autre format courant, et aucune présence de marque ou de logo.*”

Pas d'exploitation commerciale des œuvres libres de droit : “*Google s'engage à prendre garde qu'aucune copie d'ouvrage, que celle-ci soit disponible dans GoogleBooks ou sur le site de la BnF, ne soit utilisée à des fins commerciales sans l'autorisation de cette dernière.*”

Échange de bons procédés : “*Afin de pallier aux (sic) contraintes restrictives imposées à Google, chaque ouvrage numérisé de la BnF possèdera un lien redirigeant vers GoogleBooks afin d'éventuellement acheter l'ouvrage correspondant.*”

Du contexte. “*Chaque fichier devra faire l'objet d'une contextualisation, on pourra y lier un site internet où l'information aura au préalable été validée sur l'auteur, ou le courant artistique et littéraire, ou une analyse de l'oeuvre (...). Chaque fichier provenant des archives de la BnF devra être clairement identifiable par la mention BnF. Un lien hypertexte devra être mis en place qui redirigera l'internaute sur un site de “service public” qui donnera de plus amples informations et des recommandations de lecture.*”

Recréer du lien (et du service) à l'aide d'un service de type Ask a librarian tournant sur GoogleBooks et piloté par la BnF : “*Nous demandons qu'il soit ajouté, sur les pages des documents numérisés issus de la BnF, une adresse mail à laquelle les internautes pourraient envoyer leurs questions et demandes des renseignements plus précis à des personnes qualifiées. Pour cela la BnF a obtenu les financements nécessaires pour recruter deux employés à plein-temps sur ce poste.*” Malin non ? :-)

En même temps, imaginez un peu “**questionpoint**” tournant sur GoogleBooks et piloté en réseau par l'ensemble des bibliothèques contractantes ... le renouveau ou la mort du petit cheval ??

Traçabilité : “*Sur le service de Google recherche de livres, pour chaque fichier numérisé, une mention devra obligatoirement figurer pour identifier la provenance du livre. Dans la rubrique “présentation générale” vous trouverez donc la mention “issu de la BnF.”*” Important. Essentiel même.

Simultanéité de la mise à disposition : “*La copie numérique remise à la BnF devra être disponible et utilisable à des fins non-commerciales, y compris par un large public, en même temps et non pas 20 ans après celle de Google.*” Vital. Juste vital.

Transfert de compétences : “*Nous observerons les outils et la manière d'opérer de Google*

pour être capable, dans le futur, d'assurer nous-mêmes la numérisation." Après tout, pourquoi pas, on régule bien le prix du marché des avions de chasse en négociant des transferts de compétence plus ou moins larges ;-) Ceci étant, la chaîne de numérisation de Google étant un "secret industriel", la requête a peu de chances d'aboutir.

Récupérer le grisbi : *"Nous donnerons à Google une copie numérique des ouvrages que nous numériserons avec le budget obtenu par Mr Frédéric Mitterrand, en échange de quoi Google nous donnera un exemplaire numérique de tous les ouvrages francophones déjà numérisés dans Google Books."* ;-) Ah ben là j'avoue, c'est culotté. Mais les idées les plus simples – et les plus culottées – sont parfois les meilleures :-)

Et puis comme les étudiants ont de l'humour et du second degré, je ne résiste pas à l'envie de vous faire partager 2 petites pépites :

"Nous exigeons de pouvoir rencontrer les employés chargés du catalogage dans Google Books afin de pouvoir leur expliquer les notions du catalogage effectué dans les bibliothèques françaises."

Et le meilleur pour la fin :

"Il est vrai qu'il est normal d'éprouver quelques craintes concernant la pérennité et l'exploitation des fichiers numérisés mais Google possède deux programmes qui permettent une certaine garantie. AdSense qui est un moyen pour les utilisateurs de dénoncer un abus sur les droits d'auteur et HathiTrust qui assure une protection et une sauvegarde des données et des documents diffusés."

Voilà. A mon avis il y a plein d'enseignements à tirer de l'analyse fournie par les étudiant(e)s. A la fois du côté des préconisations concrètes (certaines idées et exigences pourraient aisément être reprises dans un cahier des charges au simple effort d'une légère reformulation), et aussi du côté de la perception qu'ont aujourd'hui de ce débat ceux qui demain, seront aux avant-postes. En tout cas, cela fait longtemps que je n'avais pas pris autant de plaisir à corriger des copies. Merci donc aux **A2-ML de La Roche sur Yon** et à demain en cours pour reparler de tout cela plus en détail ;-)

—

» **Article initialement publié sur Affordance.info**

» Illustration de Une par **kimdokhac** sur **Flickr**

ADRIEN

le 13 janvier 2010 - 10:54 • SIGNALER UN ABUS - PERMALINK



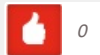
bonjour

Tout d'abord bravo à vos élèves qui ont fait part d'analyses très pertinentes. J'espère que l'élève auteur de la copie ci-dessous a eu une bonne note :) car il a très bien anticipé la nouvelle forme de partenariat que Mitterrand souhaite proposer à Google. Je dit bravo.

Récupérer le grisbi : "Nous donnerons à Google une copie numérique des ouvrages que nous numériserons avec le budget obtenu par Mr Frédéric Mitterrand, en échange de quoi Google nous donnera un exemplaire numérique de tous les ouvrages francophones déjà numérisés dans Google Books." Ah ben là j'avoue, c'est culotté. Mais les idées les plus simples – et les plus culottées – sont parfois les meilleures

Bien à vous

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

RIGGSKARA20

le 28 mai 2010 - 8:11 • SIGNALER UN ABUS - PERMALINK



According to my analysis, millions of persons on our planet receive the <http://www.lowest-rate-loans.com> from well known banks. Thence, there's great possibilities to find a credit loan in all countries.

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE