

CONTRAT GOOGLE/BIBLIOTHÈQUE DE LYON : L'OMBRE D'UN DOUTE...

LE 13 DÉCEMBRE 2009 LIONEL MAUREL (CALIMAG)

Il est étonnant de voir que deux semaines après sa révélation, le contrat liant Google à la Ville de Lyon pour la numérisation du fonds ancien de sa bibliothèque n'a pas réellement fait l'objet d'analyses détaillées d'un point de vue juridique, alors que dans le même temps le débat sur la numérisation du patrimoine continue d'occuper le devant de la scène médiatique.

Il est étonnant de voir que deux semaines après sa révélation, **le contrat liant Google à la Ville de Lyon pour la numérisation du fonds ancien de sa bibliothèque** n'a pas réellement fait l'objet d'analyses détaillées d'un point de vue juridique, alors que dans le même temps le débat sur la numérisation du patrimoine continue d'occuper le devant de la scène médiatique.

J'avais commencé à essayer de sortir les points essentiels des clauses **il y a quinze jours dans mon tour de veille hebdomadaire**, notamment pour cerner ce qui rendait ce contrat différent de ceux que nous connaissons déjà (**université de Michigan** et **université de Californie**).

Depuis, on peut lire des choses distrayantes (**comme ici chez Télérama**) dans ce registre philosophico-polémique qui tient lieu de débat public sur la question de la numérisation dans notre pays, mais des études serrées et argumentées du contrat : je n'en trouve point. On tombe aussi hélas sur des choses très approximatives, voire inexactes comme **cet article chez Actualitté**, qui titre d'une façon laissant entendre que le contrat n'imposerait aucune exclusivité (ce qui est faux).

Quand on trouve des commentaires plus précis (**comme ici chez Archimag**), ils se focalisent en général sur l'exclusivité commerciale de 25 ans consentie au profit de Google par la Ville de Lyon, aspect certes important, mais qui est loin d'être le seul méritant vigilance et attention.



J'aimerais ici revenir sur deux aspects qui me paraissent essentiels dans ce contrat et qui n'ont pour l'instant pas vraiment fait l'objet d'analyses : **la question de l'exclusivité d'indexation et celle des possibilités de réutilisation des fichiers numériques** remis à la Ville de Lyon.

Avant d'aller plus loin, je vous conseille d'aller relire le billet publié par Olivier Ertschzeid sur son blog en 2006 lorsque les contrats américains ont été révélés "**Contrat californien et eugénisme documentaire**", histoire de bien se remémorer à quel niveau d'enjeu on se situe.



[...] le marché de dupe est le suivant : chacun des 2 partenaires reçoit “sa” copie, une copie à usage interne si l’on veut. Mais le marchand (Google) s’ouvre tous les droits sur la sienne et les copies de la sienne (impression, téléchargement, revente ...) et impose au bibliothécaire un usage fermé et stérile de la sienne (pas de revente ni de cession, pas de téléchargement depuis les sites universitaires, etc.). Une forme revendiquée d’eugénisme documentaire [...] Chaque nouvelle bibliothèque contractante, en même temps qu’elle assure une visibilité de ses fonds et à l’impression de contribuer à la dissémination mondiale de la culture, fait faire un irrémédiable pas en arrière de plus à l’ambition d’une bibliothèque universelle.



C’est à cette aune que je veux examiner le contrat lyonnais en essayant de déterminer s’il contient **des éléments problématiques du point de vue de l’accès à l’information**, à savoir un déséquilibre au niveau des accès et des usages entre ce que l’on pourra faire à partir de Google Book Search et ce qui sera possible à partir des fichiers de la Bibliothèque de Lyon.

La première chose à prendre en considération de ce point de vue est **la question centrale de l’exclusivité d’indexation**. Il y a deux semaines dans mon rapide commentaire du contrat, j’indiquais que le Cahier des Clauses Techniques Particulières (CCTP) qui a été publié paraissait présenter la particularité de ne pas comporter d’exclusivité d’indexation, tout en ajoutant qu’un certain doute persistait tout de même.

Pour bien comprendre de quoi il retourne, il faut partir des deux contrats américains.

Ces derniers contiennent en effet une clause explicite qui interdit aux bibliothèques partenaires de Google de **laisser les robots des moteurs de recherche indexer le contenu des fichiers** qu’elle diffuseront à partir de leur propre site. La clause est formulée de manière identique dans le contrat de l’université de Michigan (4.4.1 p. 5) et dans celui de l’université de Californie (4.9 p. 6) :



“University shall implement technological measures (e.g. through use of the robots.txt protocol) to restrict automated access to any portion of the University Digital Copy or the portions of the University Website on which any portion of the University Digital Copy is available.”



Cette restriction signifie que **les moteurs de recherche concurrents** de Google (Bing de Microsoft, Yahoo !, Exalead, etc) n’auront pas la possibilité d’accéder à ces contenus. Certes, Google pourra de son côté les indexer (c’est même le but premier de la manœuvre !), mais il sera le seul à pouvoir le faire et c’est dans cette exclusivité que réside une forme d’atteinte à l’accès à l’information. Un seul chemin pour accéder aux contenus patrimoniaux à partir des recherches en ligne : celui que Google aura décidé de mettre en place. L’exclusivité d’indexation présente **une importance économique fondamentale**. Dans **la bataille du capitalisme cognitif** dont l’affaire Google Book est une manifestation éclatante, un des enjeux principaux consiste à savoir qui sera en mesure de mettre en place le “Portique” par lequel les utilisateurs devront passer pour accéder à l’information. De ce point de vue, pouvoir disposer d’une exclusivité d’indexation confère un avantage décisif à celui qui peut s’en prévaloir.

Lorsqu’on lit le contrat lyonnais, une des premières choses qui frappent, c’est que **cette exclusivité d’indexation ne figure pas inscrite en toutes lettres**. Aucune trace de ces deux lignes que ce soit **dans le CCTP** ou **dans l’acte d’engagement du marché**. Cela dit, je ne peux pas me départir d’un doute gênant, qui tient à la solution technique prévue pour

permettre l'accès aux copies numériques remises à la Bibliothèque de Lyon.

Le CCTP indique en effet que la mise de ces fichiers se fera “dans le cadre d’une solution hébergée propre à la Ville de Lyon (hosted solution)” à savoir “un service distant via Internet que [Google] hébergera, à ses frais, sur ses serveurs” (art.20 p. 12). Cela signifie que **c’est Google qui, dans un premier temps, construira à ses frais une bibliothèque numérique** permettant à la Bibliothèque de Lyon de diffuser ses propres copies. Or cette prestation n’est pas consentie sans contrepartie ; c’est le moins que l’on puisse dire.

On lit en effet plus loin dans le CCTP que “les fonctionnalités, le design et le contenu de ce service restent entièrement sous le contrôle du titulaire. Par design du site, il faut entendre : l’ergonomie, la présentation, l’interface et les fonctionnalités techniques de la bibliothèque numérique propre à la ville qui sera hébergée par le titulaire”.

Vous lisez comme moi : **les fonctionnalités de ce site restent “sous le contrôle” de Google**. Cela signifie que même si l’exclusivité d’indexation ne figure pas de manière explicite dans le contrat, je ne vois rien d’un point de vue juridique qui empêche Google de bloquer les robots des moteurs de recherche concurrents afin qu’ils n’indexent pas les contenus de cette *Hosted Solution*. Et symétriquement, il n’y a pas moyen contractuellement d’imposer à Google qu’il le fasse.

Un peu plus loin encore, le CCTP ajoute : “la recherche des ouvrages imprimés par l’utilisateur final doit pouvoir se faire, via Internet, aussi bien à partir du contenu des textes (full text) que des métadonnées bibliographiques”. On pourrait voir dans cette clause (les mots “full text”) une forme de garantie, mais Google aura certainement satisfait à ses obligations s’il permet simplement l’accès aux contenus de la Bibliothèque par son propre moteur.

Il existe néanmoins dans le contrat une sorte de soupape qui donne théoriquement à la Ville de Lyon les moyens de conjurer ce risque. A l’article 24 du CCTP, on peut lire : “La Ville de Lyon peut librement constituer avec les ouvrages imprimés numérisés par le titulaire, au fur et à mesure de la réalisation de la prestation, sa propre bibliothèque numérique et la rendre consultable par le public gratuitement, sur place ou via Internet”. Cela signifie qu’en plus de Google Book Search et en plus de la *Hosted Solution* développée par Google, la Ville de Lyon garde la faculté de monter une autre bibliothèque numérique (une troisième donc...), par ses propres moyens et rien alors dans le contrat n’empêche de l’ouvrir à tous les moteurs.

Garantie importante, j’en conviens...

Mais il faudra voir si concrètement la Ville de Lyon consentira à faire cet effort, alors qu’un accès aux fichiers existe déjà par le biais de la *Hosted Solution*, offerte gratuitement. Et ici encore, je vous recommande de relire ce que disait Olivier Ertzscheid en 2006 **dans son billet** :



[...] pour parachever le tout, nos chères bibliothèques n’ont d’autre choix que d’accepter le cadeau “empoisonné” de Google qui consiste à opter pour une solution hébergé, le temps de la numérisation : c’est à dire que les oeuvres seront disponibles sur le site de Google uniquement, le temps que les universités aient mis en place les capacités de stockage et de bande passante nécessaires ... Vous imaginez bien qu’une fois que vous aurez consulté une oeuvre sur Google Books (qui entre temps aura largement communiqué sur l’augmentation de l’offre disponible) vous vous empresserez (si vous êtes au courant que “ça y est, la bibliothèque a récupéré son exemplaire et le propose dans son Opac”), vous vous empresserez, disais-je, d’aller trouver le site de la bibliothèque de l’université qui possède l’ouvrage, puis de trouver sur ce site comment est-ce qu’on fait pour accéder à leur p—n d’Opac, puis de rechercher de nouveau l’ouvrage en question, puis de vous empresser de mettre ce si beau site de bibliothèque dans votre liste de favoris qui en comporte déjà 250 pour pouvoir y revenir quand bon vous semble. Hein ? Quoi ? Vous ne ferez pas ça ? Vous irez plutôt interroger directement Google Books ? Oooh, ben oui mais alors si personne ne fait d’efforts ...



Vous comprendrez que je puisse avoir l'ombre d'un doute, car si l'exclusivité d'indexation n'est pas inscrite formellement en droit, il existe un risque qu'elle puisse persister dans les faits.

A ce risque de déséquilibre dans l'accès, s'ajoutent des clauses qui vont certainement provoquer **un déséquilibre dans les usages**.

Tout se joue encore à l'article 24 du CCTP.

On y lit que *“La Ville de Lyon peut permettre le téléchargement gratuit des images numérisés d'un ouvrage à partir de sa propre bibliothèque numérique, en tout ou partie, à condition que ce soit à l'unité, pour un usage individuel”*.

“Images”... “Individuel”... deux mots qui pèsent lourd...

Il faut comprendre en effet que **les utilisateurs n'auront accès par le biais du téléchargement qu'au mode image, et non au mode texte**. Tout ceux qui ont déjà sérieusement travaillé sur un ouvrage numérisé savent que c'est une restriction très forte, notamment pour les usages d'étude et de recherche. C'est se priver d'une des principales plus-value qu'apporte la numérisation. Du mode image à la bibliothèque, tandis que dans le même temps une grande partie du contenu libre de droits de Google Book Search est déjà récupérable au format ePub... hum...

L'autre restriction drastique découle du fait qu'il **ne pourra être fait qu'un “usage individuel” des fichiers téléchargés à partir du site de la Ville de Lyon**. Est-il besoin de rappeler par exemple qu'un usage pédagogique, par définition, n'est pas un usage individuel ? Qu'il en est de même pour la plupart des usages de recherche ?

L'enseignement, la recherche comportent une dimension collective et publique qui n'est pas compatible avec un usage limité à la dimension individuelle.

De la même manière, **l'usage individuel entraîne qu'il ne sera pas possible de réutiliser ces fichiers en ligne sur Internet**. Or la réutilisation constitue un enjeu majeur de la numérisation aujourd'hui. Il fut un temps où la consultation en ligne sur le site de la bibliothèque numérique était le but ultime de la numérisation du patrimoine. C'est de moins en moins vrai aujourd'hui. Les utilisateurs ont certes besoin de consulter, mais ils veulent aussi emporter, réafficher, modifier, partager les documents, en accord avec les nouveaux usages du web. La faculté de pouvoir disséminer – et non plus seulement diffuser – les documents devient dès lors essentielle.

Or dissémination et réutilisation ne sont pas compatibles avec l' “usage individuel”.

Tout usage en ligne dépasse ce cadre étroit. L'article 24 du CCTP ajoute même pour que ce soit bien clair : *“La Ville de Lyon s'engage également à empêcher tout tiers de : [...] redistribuer toute partie de sa bibliothèque numérique”* (je serai d'ailleurs curieux de voir comment on peut satisfaire une telle exigence lorsqu'on permet par ailleurs le téléchargement).

Google a d'ailleurs bien compris l'importance de permettre aux utilisateurs de s'appropriier les documents et de les rediffuser par eux-mêmes. Il a même mis en place une solution technique pour ce faire, sous la forme d'un système de vignette exportable (*embed*), très bien fait. J'affiche par ce biais le petit extrait que vous voyez ci-dessous (choisi complètement au hasard) et ce faisant, vous conviendrez que je fais à l'évidence une utilisation du document numérique plus large que l'usage individuel. Je doute que de telles fonctionnalités puissent être mises en place par la BM de Lyon à partir de son propre site.

DES
ANCIENNES BIBLIOTHEQUES
DE LYON,
ET EN PARTICULIER DE CELLE DE LA VILLE.

LA belle situation de Lyon au confluent de deux fleuves navigables, et la fertilité de son territoire, y créèrent dès son origine un grand commerce. Celui-ci y apporta des richesses, et avec elles les jouissances du

Ajoutons que si les utilisateurs ne pourront pas disséminer les fichiers, la **Bibliothèque ne sera pas plus en mesure de le faire elle-même** puisque les partenariats sont limités par le CCTP à *“la consultation des fichiers numérisés [par Google] dans le cadre de la bibliothèque numérique propre à la BML et sans transmission des fichiers aux partenaires”*. Le contrat “fixe” les documents numériques sur le site de la BM de Lyon.

Ce qui est particulièrement dérangeant avec cette restriction à l’usage, **c’est qu’elle est plus forte que celles mises en place par les conditions d’utilisation de ... Google Book Search** ! On accède à ces conditions lorsque l’on télécharge un ouvrage et elles nous disent que l’on peut réutiliser les documents à condition de ne pas faire un usage commercial des fichiers et de ne pas supprimer l’attribution (c’est-à-dire le filigrane discret “numérisé par Google” qui figure en bas de chaque page).

Or être limité à un usage non commercial n’empêche pas de faire une utilisation pédagogique ou de recherche des fichiers, pas plus que la réutilisation en ligne sur Internet et la dissémination. A la différence de l’usage individuel qui bloque beaucoup plus de choses...

Malgré ces ombres que je discerne dans l’accord entre Google et la Ville de Lyon, je n’irais pas jusqu’à employer le terme “d’eugénisme documentaire” à son sujet. D’abord parce que je n’ai pas envie de marquer un point Godwin et d’autre part parce que ce serait manquer d’objectivité dans l’analyse de ce document.

Si on le compare avec les contrats de l’université de Michigan et de l’université de Californie, l’accord lyonnais est plus ouvert et il apporte plus de garanties. Il est très important d’insister sur ce point.

Il n’empêche que cet accord comporte toujours un risque de mise en place d’un déséquilibre des accès du point de vue de la liberté d’indexer les contenus. Et qu’il va provoquer un déséquilibre des usages à causes des restrictions imposées à la réutilisation des fichiers.

Ce qui m’étonne le plus dans cette histoire, c’est qu’on ne compte plus le nombre de personnes distinguées qui sont intervenues dans le “débat” pour se fendre d’un article à la Darnton, en montant sur de grands chevaux théoriques et nous gratifier de grandes envolées de futurologie épistémologique.

Mais pour ce qui est de se confronter directement avec les clauses du contrat... pas grand monde soudainement...

Il ne suffit pas qu’un document soit rendu public pour l’exigence démocratique soit satisfaite. Ce qui importe, c’est que cette publicité soit le point de départ de l’analyse critique.

Dans ces conditions, si j’ai l’ombre d’un doute, c’est surtout peut-être sur notre capacité à nous saisir d’une question aussi complexe que celle que nous pose Google Book Search.

Un billet publié à l’origine sur le blog S.I.Lex

MIKLOS

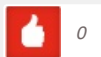
le 13 décembre 2009 - 22:19 • SIGNALER UN ABUS - PERMALINK



J'avais déjà parlé en 2005 de cette indexation exclusive que Google effectue (voir lien ci-dessus, note 12) sur ses contenus. Plus récemment, j'ai comparé un autre aspect: comment Google indexe diverses bibliothèques et portails numériques: le sien (Google Books), et les autres (Europeana, Gallica...). Les chiffres se passent de commentaires, et c'est connu depuis longtemps.

Quant à la capacité du contractant de faire indexer les fichiers que Google lui aura retournés, je ne sais pas s'ils les fournissent avec OCR intégré et indexable – ce qu'ils font fort bien (voire remarquablement bien) pour les contenus se trouvant dans Google Books.

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE

1 ping

Les tweets qui mentionnent Contrat Google/Bibliothèque de Lyon : l'ombre d'un doute... | Owni.fr -- Topsy.com le 13 décembre 2009 - 21:12

[...] Ce billet était mentionné sur Twitter par sir.chamallow et Kleiber nicolas, thierry ry. thierry ry a dit: Contrat Google/Bibliothèque de Lyon : l'ombre d'un doute... | Owni.fr <http://bit.ly/8LMds1> [...]