

Rapport Tessier sur la numérisation du patrimoine écrit

Stéphane Bortzmeyer

<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 5 février 2010

<https://www.bortzmeyer.org/rapport-tessier-num.html>

Le 12 janvier, Marc Tessier a remis au Ministère de la Culture son très intéressant rapport sur « La numérisation du patrimoine écrit <<http://www.culture.gouv.fr/mcc/content/download/3520/23115/file/Rapport%20sur%20la%20numerisation%20du%20patrimoine%20ecrit.pdf>> », rapport qui s'inscrit dans le vieil affrontement entre les propositions de Google de numériser le fonds des bibliothèques (tout en mettant la main dessus) et les projets nationaux comme Gallica, riches en proclamations martiales et pauvres en contenu accessible.

Le titulaire actuel du poste de ministre, qui avait signé la lettre de mission, est un abyssal zéro, qui s'est récemment fait connaître par des déclarations « à la chinoise » sur les dangers des blogs <http://www.lepost.fr/article/2010/01/22/1900520_frederic-mitterrand-fustige-les-blogs.html> (tous anonymes, forcément anonymes) pour l'harmonie de la société. Mais cela n'enlève rien à l'intérêt du document.

Le rapport ménage la chèvre et le chou, comme souvent dans les rapports officiels mais c'est déjà un progrès par rapport à un certain discours « astérixien » où les français ont toujours raison et où tout le mal vient des États-Unis, patrie du profit (comme si les éditeurs français n'étaient pas également des capitalistes ne pensant qu'au profit). L'ancien directeur de la BNF, Jeanneney, s'était fait une spécialité de ce discours <<http://www.ecrans.fr/Jeanneney-prefere-Gallica-a-Google,7944.html>>.

Ainsi, l'offre de Google de numérisation des bibliothèques n'est plus, dans ce rapport, diabolisée bêtement, mais analysée sérieusement, avec ses forces (les moyens matériels de Google) et ses faiblesses (un certain désintérêt pour les métadonnées, puisque Google compte surtout sur son moteur de recherche plein texte). Comme pour tout partenariat avec une entreprise capitaliste, le rapport appelle à la prudence lors des négociations et demande que celles-ci se fassent sur une base d'égalité entre partenaires. Aujourd'hui, Google approche souvent les bibliothèques comme un donateur approche des SDF, alors que Marc Tessier insiste sur le fait que chacun des partenaires a une compétence spécifique à apporter. Il demande également que les fichiers résultant de la numérisation soient distribuables par d'autres que Google, et critique les accords secrets imposés par Google.

À noter qu'une très bonne annexe concrète du rapport, l'annexe 3, faite par Alban Cerisier, rappelle les points techniques essentiels liés à la numérisation et aide à comprendre certaines questions qui pourraient paraître byzantines.

Surtout, ce qui est la première fois dans un rapport officiel, la supériorité de Google sur des offres franco-françaises comme Gallica n'est plus simplement présentée comme injuste mais aussi analysée comme le résultat des faiblesses de l'offre françaises. Les grands projets pilotés de manière ultra-bureaucratique, avec profusion d'acteurs et de réunions, puis délégation du vrai travail à une SSII n'ont pas toujours engendré une offre de qualité... Le rapport cite Gallica qui n'a toujours pas d'URL `<https://www.bortzmeyer.org/beaux-urls.html>` stables et référencables.

C'est ainsi que le rapport note à juste titre que Europeana n'est qu'une page d'accueil, une étiquette qui s'applique à plusieurs projets existants mais qui n'a aucun contenu propre. Bref, les européens ont du pain sur la planche.