

Peut-on vivre sans Google ?

La question posée maintenant est celle des risques présents dans le croisement du web mondial et de la recherche locale, sur nos ordinateurs portables.

Les mêmes outils, le moteur Google est l'un d'entre eux, permettent aujourd'hui de chercher indistinctement dans la toile mondiale et dans les tréfonds de nos disques durs.

Essayons de retracer rapidement l'évolution qui nous mène à cette situation avant d'en tirer des constats et d'en analyser les risques ou les opportunités.

CHAPITRE PREMIER : LA DERIVE DES CONTINENTS

Le web naît officiellement en 1994 avec la publication d'un article de Tim Berners Lee « L'hypertexte et le CERN ». Pour que le web existe il faut la conjonction de trois éléments distincts : des adresses (URL) permettant de localiser l'information, des navigateurs (permettant d'y accéder) et un format d'encodage (HTML) permettant d'afficher l'information récupérée. Sans oublier les protocoles d'échange de données (http et autres TCP-IP). Petit à petit cette formidable base de connaissance et d'information qu'est le web prends son « corps digital. »

Vers la fin des années 90 (**première période**) nous disposons ainsi d'un web public (le « www ») indexé par les moteurs et contenant différents types d'informations. A ses côtés, un web privé, opaque, se constitue via les sites d'entreprise (notamment) qui déploient derrière leurs pages web des bases de données d'informations dynamiques (php entre autres), invisibles à l'œil des moteurs. Tout cela cohabite avec nos habitudes de travail et le contenu de nos disques durs, « évidemment » inaccessible aux moteurs ainsi qu'avec un web « intime » celui constitué de la trame de nos messages électroniques (mail). Des ces 4 briques (web public, privé, disques durs, web intime), seule la première est accessible aux moteurs.

La question qui permet alors de scinder l'information en « visible / invisible », « indexée / non-indexée » est celle de la nature de l'information : Quoi indexer ? A laquelle on répond logiquement : « ce qui est délibérément mis en accès dans les espaces indexables du web public. »

Un pas est franchi (**deuxième période**) à l'heure actuelle avec « l'indexabilité » des 4 espaces décrits ci-dessus. Il n'est plus aucun d'entre eux pour échapper à ces formidables machines à indexer que sont les moteurs, Google en tête : le web public se déploie de manière exponentielle, le web privé ou invisible l'est de moins en moins (le premier symptôme fut l'indexation des fichiers « non-web » de type .doc ou .pdf) et dans ce même espace les entreprises emploient in interne des solutions de recherche « corporate » vendues par les mêmes acteurs (là encore,

L'avenir de la recherche d'information sur le web, peut-on vivre sans Google ?

mais il n'est pas le seul, Google), mais surtout, le contenu de nos mails bascule vers des services de webmail avec des espaces de stockage défiant l'imagination (Mailnation.net offre gratuitement 1000 Gigas de stockage ...), et last but not least, nos disques durs eux-mêmes sont désormais indexés et visibles par les moteurs. Il n'est donc a priori plus rien de notre vie intime et professionnelle qui échappe (théoriquement) à ces étranges scrutateurs. Cette part d'intimité dévoilée, indexée, est même accrue par le phénomène des blogs qui déversent sur la toile les états d'âme d'une multitude d'individualités.

La question clé pour saisir ce qui se joue à l'heure actuelle n'est donc plus de savoir « quoi indexer ? », tout étant indexable, mais de savoir « qui indexe », ou plus précisément « qui a capacité de détenir des informations suffisamment complètes sur nos vies professionnelles et intimes pour pouvoir à partir de là regrouper des bases de données hétérogènes et les utiliser à des fins commerciales ou autres. » Mais là encore la réponse est déjà posée. Google mais aussi Yahoo ! et MSN disposent d'outils et de services web qui leur permettent cela. Pour en rester sur la cas Google, Gmail collecte et scanne le contenu de nos mails, Blogger indexe nos états d'âme de blogueurs, Gdesktop vient fouiner dans nos disques durs.

Pour l'instant (et pour ce que l'on en sait) ces bases de données, ces bases d'index sont encore distinctes et liées au service qui les supporte. Mais la société mère est la même. C'est Google qui est le dépositaire de l'ensemble de ces données librement ou distraitemment confiées.

Et nous voilà à la **troisième période** de cette dérive des continents informationnels : la fusion complète et totale des services et des bases de données associées au sein d'un seul et même index. La question posée est ici celle que John Battelle appelle « la base de données des intentions ». Après le « quoi indexer » de la première période, après le « qui indexe quoi » de la seconde, il faut maintenant déterminer le « Pour quoi cette indexation ? » Et ici la réponse s'affirme comme une évidence : il s'agit de monétiser, de marchandiser des services, des biens, des échanges, des idées ... Car de l'interopérabilité entre ces différentes bases de données, Google a dès aujourd'hui la possibilité d'établir des inférences et des recoupements pour savoir qui nous voyons, ce que nous nous disons dans nos mails, avec qui nous échangeons, sur quels sites naviguons nous le plus souvent, quels livres ou quels types de biens nous achetons en ligne. Bref un ensemble d'éléments qui valent de l'or quand il s'agit d'établir des campagnes de publicité ciblées ou d'afficher des publicités contextuelles lorsque nous relevons notre courrier électronique.

L'avenir de la recherche d'information sur le web, peut-on vivre sans Google ?

Donc résumons : **la dérive des continents informationnels décrit le mouvement de ces gisements d'information qui, à l'inverse de la marche du monde, se rapprochent et fusionnent inexorablement entre les mains d'un très petit nombre de sociétés commerciales dont l'objectif avoué est de maintenir en vie un modèle économique consistant à monétiser et à marchandiser le maximum de services et de biens communs d'information possibles.**

On l'aura compris, cette dérive n'est pas sans conséquences (doux euphémisme) sur notre manière d'interagir en réseau et sur nos modes de traitement, de recherche et d'accès à l'information. Elle impacte directement notre environnement cognitif : témoin cette stupeur lorsque j'ai vu pour la première fois s'afficher sur la page de résultats de google (je venais d'installer l'outil de desktop searching) des pages web « classiques » et des documents dormant dans les entrailles de mon disque dur.

CHAPITRE SECOND : LA CONFUSION DES PRATIQUES INFORMATIONNELLES

L'un des corrolaires directs de cette étrange hybridation entre recherche locale et globale est d'entraîner une grande confusion des pratiques informationnelles. Surtout qu'à l'heure actuelle ces dernières sont loin de se résumer à une activité de « recherche » ou « d'information retrieval » comme disent les anglo-saxons : « chercher » donc mais aussi « communiquer », « partager », « annoter », « organiser », « s'orienter » sont autant de verbes d'actions engagés dans nos pratiques quotidiennes d'information. Et là encore, Google et ses deux acolytes (Yahoo ! et MSN) brouillent considérablement les cartes.

Voici une petite liste de pratiques informationnelles et des outils associés. Pour le seul Google il nous est aujourd'hui possible de :

- chercher (via le moteur Google) dans des espaces, on l'a vu auparavant, distincts.
- Communiquer de personne à personne via Google Talk
- S'orienter à l'échelle du monde (via Google Earth) ou à celle de notre quartier (via Google Local et Google Maps)
- Partager des contacts et des réseaux sociaux via Orkut
- Diffuser de l'information (intime ou non) via Blogger
- Tagger/indexer des pages web (sur le principe des folksonomies) via différentes API disponibles autour de Google
- et l'on pourrait continuer la liste ad aeternam : acheter, vendre, comparer (Froogle), correspondre (Gmail) ...

L'avenir de la recherche d'information sur le web, peut-on vivre sans Google ?

Sans oublier la plus problématique de ces pratiques informationnelles, la plus spécifique parce que la plus liée aux capacités du raisonnement humain : ORGANISER. Or il s'agit, on le sait de l'ambition et de la devise affichée de Google : « *Organiser l'information à l'échelle de la planète.* » Admettons (par manque de temps) qu'il s'agisse là d'une compétence entrant dans les capacités d'un outil reposant presque exclusivement sur des logiques statistiques de matching, et qu'il ne soit besoin d'aucune capacité intellectuelle particulière pour mettre en place une bibliothèque mondiale (Google Print) ... admettons disais-je même si l'on aura compris l'inanité de la précédente proposition ...

Or bien plus qu'une organisation des connaissances (« informations » serait déjà plus réaliste) à l'échelle de la planète c'est à une véritable dérégulation des pratiques que nous convient nos chers moteurs au travers des services (certes parfois innovants) qu'ils mettent en place et surtout au travers de la fusion de ces services entre eux au nom d'une personnalisation toujours accrue (l'exemple en la matière étant donné par Yahoo ! et l'ensemble des services regroupés sous MyWeb2.0 ou encore Yahoo!360)

Envisageons le problème selon deux axes : en bas, sur l'axe des abscisses (PERMETTRE DE), la longue liste des pratiques informationnelles précitées (« *rechercher, organiser, communiquer, échanger, tagger* » ...). A la verticale (VENDRE) , sur l'axe des ordonnées, la non moins longue liste des services marchands (ou en passe de le devenir) proposés par les moteurs : il s'agit de vendre des livres, des accès Wifi, de la téléphonie, des mots-clés payants, demain peut-être des laptop ou des environnements web pour remplacer les OS existants ...

La question qui se pose pour les trois grands (GYM : Google, Yahoo ! MSN) ou plus exactement la question que nous devons leur poser sachant que c'est l'ensemble des croisements possibles entre les deux axes qui règle leurs logiques commerciales, la question que nous devons leur poser, disais-je, est celle de savoir :

- selon quels référents et avec quelles limites ces croisements seront faits ?
- quels sont les modèles sur lesquels ils comptent s'appuyer pour mener à bien certains de ces croisements ? Songeons à la bibliothèque universelle selon Google (Print) peut-on mener au sein d'un même espace, d'une même interface une mission de mise à disposition des connaissances pour le plus grand nombre et une autre de rentabilité pour la vente d'ouvrages ? Le modèle préconisé par les adeptes de bibliothéconomie n'est pas exactement le même que celui préconisé par les intérêts boursiers de Wall Street ...
- Quelle est la logique de ces sociétés : « donner accès » ou « prendre (des) parts » ??

L'avenir de la recherche d'information sur le web, peut-on vivre sans Google ?

- Enfin quels croisements seront faits entre quels services pour éviter la mise en place de ces « *bases de données des intentions* » qui mettent en jeu et en cause notre droit à « s'informer en paix » sans avoir à chaque instant de questions à se poser sur « comment et par qui seront réutilisées les informations que je mets en ligne. »

Alors peut-on vivre sans Google ? Bien évidemment.

Peut-on s'informer sans Google ? Aujourd'hui oui. Mais demain ?

Comme l'a sûrement rappelé Jean Véronis, les pires scénarios Orwelliens sont imaginables. Plus nous serons à en débattre, plus nous serons à affirmer la nécessité absolue d'une vraie pédagogie de l'information (et des outils d'information) à l'école et à l'université, plus nous serons à réfléchir ensemble aux alternatives possibles en évitant de sombrer dans un discours catastrophiste oublieux des extraordinaires avancées que ces outils ont permis pour la démocratisation et l'accès à l'information, et plus lesdits scénarios orwelliens auront de chance de se dissiper.

Une chose m'importe encore : rappeler qu'aujourd'hui quand nous évoquons la question des outils de recherche (Google ou un autre mais Google en particulier) nous ne devons pas nous tromper d'échelle : nous ne parlons pas d'informatique et il ne s'agit plus de technologie. Et il n'est plus acceptable d'affirmer (comme le fit encore il y a peu Eric Schmitt, CEO de Google interrogé sur la part de Google dans la censure des sites accessibles depuis la Chine) que « les technologies sont neutres ». Les technologies ne valent que par leur USAGE : l'histoire y compris la plus tragique nous l'a suffisamment montré. A nous (et à d'autres) d'allier pédagogie et réflexion politique sur ce que nous voulons faire demain d'un monde d'informations pour une économie de la connaissance.

Le chapitre troisième reste donc à écrire. Au delà des percées technologiques qui guideront ses pas (web sémantique) il ne pourra pas faire l'économie d'une réflexion politique sur les enjeux de la maîtrise individuelle de l'information pour la co-construction d'une connaissance collective.

L'avenir de la recherche d'information sur le web, peut-on vivre sans Google ?