

CHRISTIAN FAURÉ

La lecture dans les nuages : quelques éléments d'architecture

Je vais tenter de rejouer la distinction marxienne entre infrastructure et superstructure, au sens où les infrastructures surdéterminent les superstructures. Je suis en effet convaincu que l'infrastructure du numérique en réseau surdétermine les métamorphoses numériques du livre.

Architecture du numérique en réseau - De quel nuage parle-t-on ?
Derrière l'expression "lecture dans les nuages", il y a l'ambition de tremper le livre dans ce milieu technologique du numérique réticulaire. Comment l'informatique dans les nuages (ou sur internet) – le *cloud computing** en anglais – métamorphose-t-elle le livre, ses pratiques et son économie ?

Ce qui caractérise l'architecture du numérique c'est l'architecture client/serveur, qui n'est plus l'architecture émetteur/récepteur de type *broadcasting** du milieu technologique analogique. Le client envoie ses requêtes à un serveur, et ce dernier envoie ses réponses au client ; il y a des protocoles de dialogue bidirectionnels.

Les protocoles client/serveur ont évolué au cours des cinquante dernières années. Dans cette évolution, je retiens quatre phases liées à des entreprises emblématiques : le *hard ware* (les machines) avec IBM, le *soft ware* (les logiciels) avec Microsoft, le *net ware* (le réseau) avec Sun, le *data ware* (les données) avec Google.

C'est grâce à l'évolution des protocoles de transfert que le *cloud computing* – qui est une tendance centripète dans l'environnement distribué du web – émerge en proposant des plateformes informatiques de serveurs qui abritent une seule instance applicative. Cette mono-instanciation explique à elle seule de nombreux phénomènes :

- Ce n'est plus le logiciel qui est vendu mais son utilisation qui est louée (on achète un droit d'accès). On sort ainsi du débat logiciel libre/logiciel propriétaire puisque le logiciel ne fait plus l'objet d'une transaction.
- Rapidité d'innovation et d'évolution de la plateforme (mises à jour quotidiennes).
- Automatisation de l'accès : interface web (sans interlocuteur humain).
- Concentration des données et vision panoptique des usages.

Les technologies et les industries de transfert

Google, Facebook, Amazon sont des industriels du transfert qui n'existaient pas – ou à peine – il y a une douzaine d'années.

Quelle distinction entre transfert et transport ? Si les deux relèvent de la mobilité et du déplacement, les réseaux de transport déplacent des objets (et plus généralement de la matière), tandis que les réseaux de transfert déplacent des représentations et des symboles (et plus généralement du signifiant, dont les données sont la plus petite unité). En informatique, on transporte des bits mais on transfère des représentations.

Dans les technologies de transfert, il y a des protocoles de dialogue signifiant relatifs aux informations et données échangées entre les machines, alors que les technologies de transport ne font qu'effectuer des déplacements d'un point A à un point B. Il ne faut pas pour autant opposer réseaux de transfert et réseaux de transport, car aucun réseau de transfert n'existe sans un réseau de transport sur lequel s'appuyer. En informatique, cette articulation entre les différents protocoles et réseaux est illustrée par le modèle en couche, selon lequel ce sont les protocoles qui "changent la donne" : SMTP (pour envoyer des mails), FTP (pour échanger des fichiers), HTTP (pour naviguer sur le web).



Architecte principal Cap Gemini.
Ingénieur et philosophe.

Après des études de philosophie à l'université de Toulouse le Mirail et des études à l'École nationale d'Ingénieurs de Tarbes, Christian Faure travaille dans de grands groupes industriels (Motorola, EADS) puis dans des cabinets de conseil en organisation et en technologies des systèmes d'information (Unilog, Logica, Atos Origin). Membre du conseil d'administration d'Ars Industrialis, il axe ses recherches sur les infrastructures du numérique, en étudiant les liens entre réseaux de transports (comment accède-t-on à l'internet ?) et réseaux de transferts (nature des données transportées).

Publications

Pour en finir avec la mécroissance. Quelques réflexions d'Ars Industrialis, avec Bernard Stiegler et Alain Giffard, Flammarion, 2009

Site

www.christian-faure.net

Il y a quelques années, on a qualifié le numérique en réseau de "virtuel". Ce que l'on ne voyait pas, entre les industries du transport et du logiciel, c'est précisément les infrastructures de transfert, celles dont l'usine moderne est le *data center*. Pendant qu'on nous racontait la fable de l'immatériel, Google installait dans le monde entier plus d'un million de serveurs répartis dans 30 *data centers*. Ne pas faire la distinction entre infrastructures de transport et infrastructures de transfert, c'est ne pas comprendre la guerre de tranchées que ces industries se font depuis plusieurs années. C'est également la question de la neutralité du net.

Le transfert du propre

Par ailleurs, le *Littre* nous rappelle que "transfert" est un terme d'origine financière et juridique : on transfère des droits, des propriétés, des actions, des marchandises... Le transfert est donc par définition le domaine du propre, et pose des questions de droit et de propriété. Bien sûr, les questions juridiques sont en première ligne de la mutation induite par les industries de transferts numériques, qui émergent et s'autonomisent peu à peu des acteurs du transport.

Or que font les industries du transfert numérique ? Elles ont tendance à fonctionner sur la base du transfert de nos propres données, de nos traces numériques. D'ailleurs, la première chose que nous faisons en accédant à un service web, c'est d'accepter les "conditions générales d'utilisation", en vertu desquelles nous transférons des droits d'usage sur notre propriété numérique. Malheureusement, dans l'économie que nous proposons les industriels des réseaux numériques de transferts, il y a toujours le risque qu'en s'appropriant les données des utilisateurs, c'est-à-dire leur propriété numérique, ces derniers ne deviennent purement et simplement des "propres à rien", quand cette logique d'exploitation des données vire à la dépossession.

Identifiants et identités numériques

Avec les réseaux de transfert numérique il faut s'identifier. C'est ce que Jeremy Rifkin* a appelé "l'âge de l'accès". Contrairement aux réseaux analogiques où le maillage du territoire se faisait par région, vallée ou commune, sans identification, le maillage pour le numérique se fait au niveau des familles, des foyers et des individus.

Autour de l'identifiant numérique s'agrègent des graphes de données, le tout formant des identités numériques, utilisées pour calibrer le filet du service web et être revendues au marketing. Ce phénomène de captation des données par transferts numériques se fait bien souvent aux dépens des utilisateurs, dans la confusion et l'offuscation. Le réseau numérique est fallacieux dans le sens où il prolétarise (perte de savoirs) les comportements et participe d'une tendance à la dépossession du numérique.

Qu'en est-il de la métamorphose numérique du livre eu égard au *cloud computing* ?

Plongeons le livre dans l'environnement et le cadre que je viens de tracer : l'architecture du numérique en réseau, le *cloud computing*, les technologies de transfert, la confusion et l'offuscation.

Avec les architectures techniques actuelles, les œuvres sont hébergées dans les *data centers* d'une entreprise et, à l'autre bout du tuyau, se trouvent des lecteurs numériques (téléphones, tablettes et liseuses). Tout cela n'est possible que grâce à la connexion au réseau.

Il s'instaure alors une asymétrie des données : un acteur central avec une vue panoptique des comportements et, en bout de chaîne, un lecteur qui ne voit rien, qui reste au fond de la caverne de Platon. Nous ne voyons pas – en tout cas pas tout de suite – que nombre de données et métadonnées nous appartenant sont captées : achats, parcours dans le catalogue, recherches, avis et consultation de critiques et, de façon plus inédite, comportements de lecture (*reader analytics*). L'anonymat et la confidentialité du lecteur sont donc fortement menacés. Il faut savoir qu'un gouvernement peut demander toutes les données relatives à un individu, notamment un chef d'État. C'est un enjeu géopolitique.

“L'anonymat et la confidentialité du lecteur sont fortement menacés.”

L'existence de ces données n'est pas un mal en soi, mais il faut rester lucide sur les modèles "d'affaires afférent" qui se développent, notamment quand ceux-ci ne reposent plus sur la vente de livres mais de données comportementales sur les pratiques de lecture. Je crois que c'est ça qui choque le plus ceux qui sont attachés au livre : ces situations où le livre devient un pré-texte. Ce qui intéresse Amazon, ce ne sont pas les livres mais les traces de lecture.

Fortes de ce trésor de données, les plateformes de lecture numérique réticulaire peuvent affûter leurs techniques de profilage. Je fais ici de la prospective : vendre des conseils aux éditeurs dans le cadre de leur politique d'édition, donner des conseils d'écriture aux auteurs en leur livrant/facturant les *best practices*, etc. Ce n'est donc pas uniquement le lecteur qui risque d'être victime de la logique de dépossession évoquée précédemment, ce sont tous les professionnels du livre. Les éditeurs ne pourront pas échapper à la tentation de jeter un œil aux données amassées par les plateformes de lecture numérique, véritable graal du marketing. Les libraires et leur conseil feront pâle figure face aux suggestions algorithmisées et, si dans le meilleur cas leurs suggestions sont entendues, l'achat se fera quand même en ligne.

Sur deux modèles de lecture

Les livres et les lecteurs sont nombreux et variés, tout comme les modalités de lecture : la métamorphose du livre de poésie dans le numérique ne sera probablement pas la même que celle du roman, de l'essai ou du livre éducatif. Il y aura

vraisemblablement un feu d'artifice de trajectoires numériques. Il y a plusieurs rapports à la lecture, différents modes de lecture. Je retiens, pour les besoins de mon propos, deux types de lecture : la lecture extensive (appropriée au roman) où l'on constate une autonomie du texte qui se suffit à lui-même, et la lecture appareillée, qui s'accompagne d'une activité s'appuyant sur des techniques d'annotation et sur des "appareils critiques".

C'est souvent au sujet de la lecture extensive qu'il y a des dissensions, des tensions et des polémiques. C'est une bonne chose que ces questions liées au livre numérique soient débattues, mais il y a une équivoque dans ce débat si l'on s'en tient au mode de la lecture extensive.

Pour moi, ce n'est pas dans cette modalité de lecture que le livre a fait défaut en tant que support. Ma thèse, discutable, est que les enjeux de la lecture dans les nuages passent d'abord par la lecture appareillée, et c'est la raison pour laquelle les études numériques – les *Digital Studies* – sont si précieuses.

Les *Digital Humanities*

Les *Digital Studies* désignent plusieurs types de recherches : les *Digital Humanities*, les *Cultural Analytics*, les *Software Studies*... À chaque fois, c'est une logique d'interdisciplinarité mêlant disciplines littéraires et scientifiques.

L'origine des *Digital Humanities* se situe dans la rencontre, à la fin des années 1940, entre le père Roberto Busa (un jésuite italien féru de Saint-Thomas d'Aquin) et Thomas J. Watson (l'un des fondateurs d'IBM), donc entre la scolastique et IBM. L'œuvre de Roberto Busa est un index – l'*Index Thomisticus* – de tous les termes présents dans les œuvres de Saint-Thomas et de ses commentateurs, un appareil critique disponible sur le web qui rassemble aujourd'hui plus de 22 millions d'entrées. Il a fait appel à IBM pour automatiser la génération des entrées de cet index.

Toutes les techniques mises au point par Roberto Busa correspondent au cahier des charges réalisé plus de 50 ans plus tard par Google avec son moteur de recherche. Les Humanités entrent dans le champ scientifique avec cette idée révolutionnaire de faire du calcul sur des textes écrits pour automatiser la constitution des appareils critiques (table des matières, notes, sommaire, glossaire, index, etc.).

Les *Cultural Analytics*

Si les *Digital Humanities* correspondent aux appareils critiques numérisés, les *Cultural Analytics* correspondent à la numérisation des pratiques de lecture.

La compréhension de "culture" par les *Cultural Analytics* est très ouverte, elle va jusqu'à inclure toutes les marchandises et les services de ce marché culturel (par exemple l'*entertainment*) qui déborde donc des questions de l'économie du livre. Mais ce sont les analyses des pratiques culturelles qui sont visées. Les *Analytics*, eux, font référence aux tableaux et schémas de mise en forme de nombreuses données quantitatives utilisées dans le

commerce ou la science. Ce sont les mêmes *Analytics* que l'on retrouve dans les statistiques de consultation d'un site web. Chez Lev Manovich*, l'accent est clairement mis sur la visualisation des grands corpus de données, dans une logique de flux souvent associée aux données en temps réel.

Il y a le rêve affirmé de pouvoir voir et surveiller (au sens de *monitoring*) les flux d'échange, de production et de consommation de l'ensemble des données et métadonnées culturelles. On comprend que, pour une maison d'édition, ces données permettent d'avoir une réactivité sur le marché en sortant des livres qui marchent à un moment donné. La dérive potentielle, c'est une logique d'édition qui ne produise que des clones et finisse par s'appauvrir. L'aboutissement est toujours la visualisation d'un vaste corpus de données, ces *Readers Analytics*, que cherchent à développer ceux qui veulent promouvoir la lecture dans les nuages.

Pharmacologie positive

En résumé, ce sont les appareils critiques du livre qui ont préparé de longue date la numérisation du livre. Du côté de la lecture appareillée, il y a une forte continuité dans ce passage au numérique, au sens où le numérique ne fait qu'automatiser ce qui existait déjà, mais les appareils critiques numériques restent largement perfectibles. Une véritable crise de confiance s'instaure en ce moment suite à des mésaventures d'utilisateurs de Kindle qui, après la mise à jour d'une œuvre qu'ils avaient achetée, lue, commentée et annotée, ont perdu l'ensemble des métadonnées de leur lecture appareillée.

La rupture ne réside pas tant dans la lecture numérique au sens des appareils critiques numérisés, que dans l'extimisation de la lecture (pour reprendre l'extimité de Lacan, cette intimité extériorisée). Avec le numérique, l'intimité de la lecture se retrouve exposée, notamment via le *cloud computing*.

Le message que je souhaite faire passer aux amoureux de la lecture, c'est qu'il faut s'investir dans les *Digital Humanities*, afin que les valeurs culturelles généralement associées au monde du livre puissent accompagner et influencer sur la constitution d'une politique industrielle de ces technologies de l'esprit : chacun peut et doit y contribuer, pour que les circuits longs de la lecture ne disparaissent pas dans les brumes du *cloud*. Les informaticiens n'avanceront pas sans votre contribution. Aujourd'hui, l'enjeu est éminemment transdisciplinaire.

