

Vérité et mensonge sur l'internet

Eric Guichard – Ens-Enssib-CIPh

Novembre 2015

Résumé

La question de la présence de vérités ou de mensonges en ligne est importante. Mais il nous semble instructif de comprendre les raisons pour lesquelles la rigueur argumentative est fragilisée par l'internet actuel. Les discours partisans ou dénigreur nous empêchent de comprendre ce qu'est une technique et ne laissent pas la place à une critique raisonnée de l'internet. Les moteurs de recherche nous mentent quant au nombre de réponses qu'ils prétendent nous proposer et favorisent les plus banales. Enfin, les atomes d'écriture proposés par le consortium Unicode sont incohérents et privilégient les mentions islamiques. Nous découvrons alors que les grands industriels de l'internet, alliés avec des États comme le Sultanat d'Oman, sont les premiers à décider des formes de l'écriture numérique et que les instruments qu'ils déploient désorientent plus les internautes qu'ils ne les aident. Ce sont les authentiques détenteurs du pouvoir de l'écrit, et donc les principaux façonneurs de la culture numérique. Les personnes susceptibles de contribuer à l'élaboration d'une culture numérique qui puisse être partagée par le plus grand nombre et qui puisse servir la rationalité, la pensée critique et l'érudition sont rares. Il s'ensuit un réel problème politique qui relance le débat sur la reconstruction d'une éducation populaire pertinente et bien ciblée.

1 Présentation du problème

Nos représentations et nos discours sur l'internet oscillent souvent entre deux positions trop extrêmes pour être satisfaisantes : il nous apporterait le paradis ou l'enfer.

1.1 Un internet enchanteur par nature ?

L'internet est souvent décrit comme chargé de positivité : vecteur de connaissance, de «changements de paradigmes»¹ ou de démocratie. Ainsi certains journalistes évoquaient les «révolutions facebook» du monde arabe. Nous sommes invités à croire que les techniques nouvelles induisent un monde meilleur, ce que contredit notre expérience des dernières décennies : écoles et hôpitaux aux budgets réduits, accroissement du chômage, de la précarité et de la pauvreté², multiplication des conflits guerriers.

1. Cf. le film *Us Now* qui explique comment, grâce au web collaboratif, «nous» pouvons tous nous organiser sans services publics et sans gouvernements — systématiquement présentés comme inutiles, hiérarchiques, bureaucratiques (Guichard, 2010).

2. 14 % des ménages, soit 9 millions de personnes, vivent en deça du seuil de pauvreté : http://www.lemonde.fr/economie-francaise/article/2015/12/22/en-france-la-pauvrete-a-legerement-augmente-en-2014_4836778_1656968.html, Le Monde en ligne, 23 décembre 2015.

Les historiens des techniques nous rappellent le caractère idéologique de tels discours : une «rhétorique révolutionnaire, orientée vers l'avenir est devenue le mode dominant de présentation des techniques» (Edgerton, 1998). Ils insistent sur le fait que ces cadres discursifs ne changent pas au fil du temps : seul l'assaisonnement subit quelques variantes : il y a 30 ans, c'était aux «nouvelles technologies» (ordinateur, CD-Rom) de nous garantir un futur meilleur. Ensuite cette fonction fut prise en charge par l'internet. Après l'explosion de sa bulle en 1999, nous fûmes incités à croire en les vertus magiques du web 2.0. Quand nous découvriâmes que cette formulation renvoyait essentiellement à une exploitation gratuite de l'internaute (Matthews, 2010), l'industrie de l'internet nous promit les réseaux sociaux, vite débordés de publicité ciblée. C'est aujourd'hui le tour du *cloud* de relayer cette utopie.

Aussi devons-nous rester dubitatifs face aux annonces enchanteresses qui accompagnent l'évolution de l'internet.

1.2 Un internet maléfique par essence ?

En France, les partisans du Minitel ont très tôt critiqué l'internet, supposé concurrencer la petite boîte télématique de France-Télécom. Aussi l'internet fut-il accusé de nombreux défauts : lieu de pédophiles, non sécurisé, média non régulé (Guichard, 1996)... Cette tradition s'est maintenue puisqu'en 2005, le président de la Bibliothèque nationale de France écrit un pamphlet contre Google, supposé menacer la culture française (Jeanneney, 2005). Pour autant, les solutions franco-européennes qu'il a contribuées à mettre sur pied n'ont pas fonctionné (Moatti, 2015). Étrangement, ces réactions nationalistes ou plus largement technophobes n'ont pas produit une critique de l'internet qui soit efficace ou rationnelle. Par exemple, les usages des traces de l'internet à des fins de surveillance ou de publicité³ n'ont pas fait débat avant les révélations d'Assange et de Snowden : les contempteurs de l'internet et de la technique, qui auraient pu forger des arguments contre de tels abus, n'ont pas vu la menace.

1.3 Retrouver le régime de la raison et de la critique

Cette première opposition entre société rêvée et espace totalitaire, entre promesses du futur et contraction de la liberté individuelle stimule notre questionnement sur le thème de la vérité et du mensonge sur l'internet : faut-il donner des bons et des mauvais points à l'internet en tant que nouveau média, en parant la technique de qualités et défauts intrinsèques ? La chose n'est pas réaliste si nous pensons qu'une technique atteint son régime de maturité au moment où elle a été suffisamment transformée par les humains pour ne plus avoir de rapport avec ses premières caractéristiques. Pour le dire autrement, une technique est majoritairement une production sociale — ce qui, accessoirement, contredit les partisans du déterminisme technique, qui supposent qu'elle puisse transformer une société. Il est plus intéressant de comprendre en quoi l'internet brouille nos catégories de l'entendement, rebat les cartes de la raison et de l'argumentation. Il est important de repérer les acteurs principaux de ce brouillage et de préciser leurs parts de responsabilités, sans pour autant nous dédouaner des nôtres. Par exemple, les dangers de l'internet sont connus de nous : la loi sur le renseignement a été publiquement débattue et votée en 2015. Nous pouvons, comme

3. Nous savons désormais à quel point nos goûts, nos désirs d'achat sont orientés par des algorithmes (Chavaliaris, 2012).

l'auteur de ces lignes, en signaler les dangers⁴. Nous ne pouvons plus (du moins en 2016) prétendre que ces formes de cibrage et de surveillance soient masquées sous des propos mensongers.

Les vrais mensonges sont aussi ceux que nous voulons entendre.

1.4 Descartes internautes ?

Pourquoi notre esprit critique est-il émoussé ? Deux raisons nous viennent à l'esprit. La première est que nous avons une réelle difficulté à construire des discours cohérents sur l'internet, sur la technique. Les analyses héritées des XVIII^e et XIX^e siècles sont moins efficaces que nous le pensons, en partie du fait que la technique est aujourd'hui moins objective, plus intime et plus intellectuelle qu'au temps des machines-outils. Jack Goody, sur lequel je reviendrai, définit l'écriture comme une «technologie de l'intellect». Cela signifie aussi qu'il s'écarte d'une doxa philosophique qui promouvait une pensée pure, spirituelle, intérieure.

La seconde raison est concrète. Notre écriture contemporaine nous offre de grands avantages, par exemple aux plans épistémologiques et méthodologiques. Mais elle est aussi particulièrement instable. Avant de détailler ce point, faisons un constat : l'internet actuel est compliqué. Avant toute activité «intelligente», il nous faut maîtriser une gestuelle confuse et changeante (d'une machine à l'autre, d'une année à l'autre) : sur quelle icône devons-nous cliquer pour lancer un logiciel, pour supprimer un fichier, interrompre un processus ? Ces choses ne vont pas de soi. Si nous sommes familiarisés avec les notions de bureau, d'application, de dossier, c'est aussi que nous avons été conditionnés depuis une trentaine d'années pour intégrer les catégories que l'industrie informatique a créées pour le *vulgus internautum*, qui n'est pour elle qu'un client potentiel. Descartes, Einstein ou Foucault, dont nous pouvons supposer qu'ils étaient d'une grande intelligence, ne sauraient certainement pas comment lancer un navigateur et quel parti en tirer s'ils revenaient en 2015. En revanche, ils sauraient rapidement réaliser deux types d'action :

- Distinguer, dans leurs domaines de compétence, le vrai du faux. Ils auraient vite trouvé des exemples de présences de la preuve et de la vérité sur l'internet en consultant des sites de chercheurs, de professeurs de lycée, et apprécié (pour les deux premiers) l'existence du site d'articles scientifiques en ligne <http://arxiv.org>. Tous auraient pu repérer des manifestations de mensonge ou de mauvaise foi en des lieux incitant à la confiance⁵.
- Ils auraient certainement, et peut-être mieux que nous, su théoriser et critiquer les cheminements de l'accès à la connaissance tels qu'ils sont formatés aujourd'hui par les industriels de l'internet, les publicitaires et les promoteurs naïfs de la nouveauté.

4. Cf. la «Lettre ouverte aux membres du Conseil constitutionnel», publiée dans les journaux Le Monde, Médiapart, Rue 89 et Nextimpact, accessible à divers URL dont <http://rue89.nouvelobs.com/2015/07/21/renseignement-loi-menace-lequilibre-democratie-260367>.

5. Comme la liste Theuth (theuth@listes.univ-rennes1.fr), dédiée à l'échange entre épistémologues, qui fut polluée le 8 novembre 2015 par des promoteurs des frères Bogdanov.

2 Une documentation peu crédible

Nous pouvons nous réjouir de stocker et de consulter des milliers de livres ou d'ouvrages sur des objets extraordinairement légers. Nos laboratoires photographiques et nos bibliothèques se portent désormais dans un sac, et nos facilités d'accès à ces documents sont décuplées, comme celles de mise en correspondance et de modification. En même temps, nous sommes étouffés par les centaines de formats d'images et les dizaines de formats de fichiers textuels : nous ne comprenons pas trop pourquoi nous devons disposer de logiciels différents pour lire des articles aux formats epub, pdf, odt, etc. même si nous commençons à nous familiariser avec des logiciels pivots comme `calibre` ou `libreoffice`.

2.1 Des résultats impossibles à vérifier

Notre rapport à l'informatique en réseau est plus complexe : confrontés à des milliards de pages web dont nous ne connaissons pas l'existence, fréquemment dynamiques⁶, nous faisons de plus en plus confiance en les moteurs de recherche. Ces derniers proposent à nos questions des résultats préalablement triés par le biais d'un algorithme qui met en avant les plus pertinents. Par exemple, pour la requête *entretiens d'Auxerre*, un célèbre moteur de recherche nous affirme avoir trouvé (le 23 décembre 2015) 542 000 résultats : cf. figure 1.

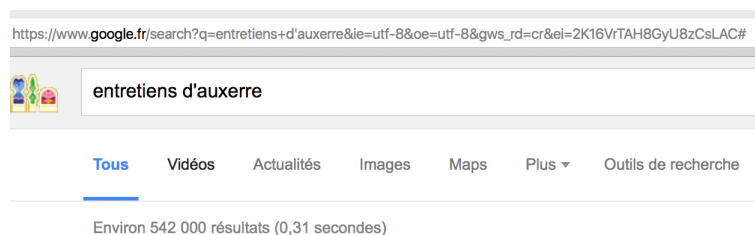


FIGURE 1 – La réponse optimiste de Google à la requête *Entretiens d'Auxerre* : 542 000 résultats. Recherche réalisée en décembre 2015.

Nous pouvons explorer les premières pages de réponses, mais dans les faits, il nous est interdit de consulter les réponses de rang supérieur à 1000. En effet, la saisie de l'adresse `https://www.google.fr/search?q=entretiens+d'auxerre&start=1000` conduit à la réponse lapidaire «Google ne renvoie pas plus de 1000 résultats par requête. Vous avez demandé l'affichage des résultats à partir du numéro 1000». Cf. figure 2.

6. Pages fabriquées par des robots en fonction de nos demandes, comme les sites qui nous donnent l'heure de départ du prochain train.

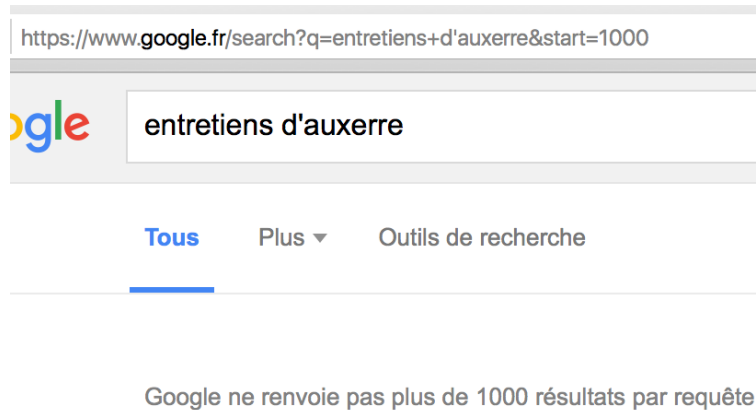


FIGURE 2 – L'interdiction de consulter les réponses de rang supérieur à 1000. Recherche plusieurs fois réalisée en décembre 2015.

Cela signifie *a priori* que nous n'aurions accès qu'à l'écume des réponses admissibles : 0,2% d'entre elles. Quelle confiance pouvons-nous avoir en une industrie privée qui nous affirme avoir trouvé une grande masse de résultats et qui refuse que nous vérifions ses dires ? La chose est d'autant plus problématique que, si nous regardons les réponses de rang 990 à 1000 (<https://www.google.fr/search?q=entretiens+d'auxerre&start=990>), nous obtenons la réponse suivante : «Afin d'afficher les résultats les plus pertinents, nous avons omis quelques entrées qui sont très similaires aux 376 entrées actuelles.». En fait, les 542 000 résultats promis se réduisent comme peau de chagrin à 376. Cf. figure 3. Notre entreprise est ici prise en flagrant délit de mensonge.

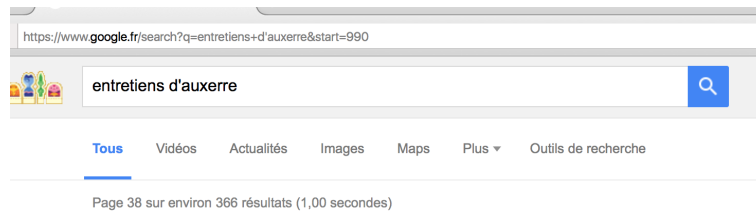


FIGURE 3 – En demandant les réponses de rang supérieur à 990, nous découvrons que Google n'en propose que 366. C'est bien moins que les 542 000 promises. Cet exemple est aisément reproductible avec tout autre requête («entretien d'embauche», etc.).

Nous pouvons le vérifier «à la main», en consultant les unes après les autres les pages de réponse. Cf. figure 4.

Afin d'afficher les résultats les plus pertinents, nous avons omis quelques entrées qui sont très similaires aux 328 entrées actuelles. Si vous le souhaitez, vous pouvez [relancer la recherche pour inclure les résultats omis](#).

Recherches associées à entretiens d'auxerre

[entretiens d'auxerre 2015](#)

[cercle condorcet auxerre](#)



FIGURE 4 – En consultant le 3 janvier 2016 les réponses page à page, nous confirmons la découverte précédente. Ce jour, Google prétend avoir trouvé «557 000 résultats» alors qu’il n’en offre que 328.

2.2 Désorienter

À ce fait, récemment découvert par l’auteur, s’ajoutent deux logiques, la première connue depuis quelques années par les experts, la seconde vérifiée tous les jours par le grand public.

Google et d’autres moteurs optimisent leurs réponses en fonction de nos centres d’intérêt (nos «profils»). De ce fait, deux personnes adressant une même requête recevront des résultats différents⁷. Avec les moteurs de recherche dont le modèle économique est fondé sur la publicité ciblée, disparaît l’espoir d’un débat scientifique serein, fondé sur l’assurance de travailler à partir des mêmes sources et références.

L’autre fait nous gêne au quotidien. Quand nous restons dans nos domaines de compétence, il nous est aisé de complexifier une requête pour trouver une réponse pertinente à nos préoccupations. En revanche, quand nous désirons consolider notre culture sur un thème moins maîtrisé, nous pouvons nous retrouver désorientés par une masse d’informations inutiles. Je reprends ici trois exemples déjà évoqués dans un précédent article (Guichard, 2014).

— Trévoux est une petite ville au Nord de Lyon connue pour son dictionnaire⁸, son parlement accordé par François I^{er} (privilège toujours refusé à la ville de Lyon) et aussi pour le monopole qui lui a été attribué en matière de fabrication des fils d’or et d’argent. Si nous nous intéressons à l’histoire industrielle de la ville et adressons une requête du type *histoire tréfilage Trévoux* nous trouvons assez rapidement des informations sur ce

7. La chose n’est pas toujours aisée à mettre en évidence. Nous la vérifions partiellement quand nous adressons une même requête de lieux différents : l’accompagnement publicitaire des réponses à nos requêtes s’adapte à notre localisation, par exemple selon que nous soyons à Paris, à Auxerre ou à Lyon.

8. Dont la première édition date de 1704. Cf. https://fr.wikipedia.org/wiki/Dictionnaire_de_Trévoux.

point. La question plus simple *tréfilerie Trévoux* donne en revanche beaucoup de propositions pour louer ou acheter un bien immobilier rue de la Tréfilerie à Trévoux, mais aucune sur l'histoire de la ville

- La recherche d'informations sur le logiciel `convert`, outil performant de traitement de l'image, induira l'internaute en erreur s'il oublie d'intégrer le mot clé «ImageMagick» (logiciel dont `convert` est le moteur). Il lui faudra faire défiler plusieurs pages avant de trouver une réponse pertinente. Des réponses supposées instructives (dotées des bons mots clés) renvoient de fait sur des sites écrits en japonais, sur des sites marocains à vendre, etc. : les sites les plus aptes à le désorienter sont les premiers présentés.
- Au début des années 2000, une solution pour se sortir d'un mauvais pas informatique (bug, erreur de syntaxe, etc.) consistait à transmettre sous forme de requête à un moteur de recherche les quelques informations arides transmises par l'ordinateur. Aujourd'hui, cette solution ne fonctionne plus : les sites de *geeks* grand public sont mis en avant par les moteurs de recherche et contiennent ces mêmes questions, posées par des internautes, sans les réponses.

Inaccessibilité de la majorité des sources, fausseté des informations quant à leur nombre, impossibilité de les comparer, désorientations multiples consécutives à la prime accordée aux sites de piètre valeur informationnelle qualifient désormais la recherche sur l'internet.

L'accès à la connaissance est-il vraiment universel ou réservé aux personnes capables d'énoncer des requêtes très précises et de résister à la surcharge cognitive induite par la prolifération de réponses banales et sources de leurres ?

3 Une nouvelle tour de Babel ?

3.1 La spécificité de l'internet

Nous pouvons prendre conscience de la différence entre l'internet et d'autres techniques en étudiant nos pratiques les plus banales : nous passons notre temps à lire (des mails, des pages web), à écrire (des requêtes, des mails, des sites web, des lignes sur les réseaux sociaux). Ces activités n'ont aucun rapport avec nos usages de l'automobile ou du lave-vaisselle. Elles sont en rapport étroit avec l'écriture au sens large : lire, écrire, compter, dessiner. La maîtrise de l'écriture conjugue ces savoir-faire, tous à dominante technique avec une réflexion sur ces pratiques et leur histoire.

3.2 L'écriture

Pour Jack Goody, l'écriture est une technologie de l'intellect (Goody, 1994), qui de plus est réflexive : une technique qui se laisse mettre en perspective, qui devient objet de réflexion. Une autre façon de préciser cette réflexivité consiste à remarquer que nous pouvons expliquer ce qu'est l'écriture avec l'écriture. Une autre technique intellectuelle dispose de cette capacité : le langage. En effet, nous pouvons expliquer ce qu'il produit et ses fonctions avec les mots de la langue, et ce même dans les sociétés sans écriture. Aussi l'écriture sollicite-t-elle un système de signes bien dosé. Ni panneaux de circulation, avec lesquels nous pouvons faire des rébus, mais pas écrire la phrase *nous sommes des panneaux de circulation*, ni totalité des signes, qui intégrerait ces fameux panneaux, les tremblements d'un alcoolique, les taches rouges sur ses mains, etc. et l'alphabet (Eco, 1988).

L'écriture est en fait une technique qui dysfonctionne : elle traduit mal l'oralité (le langage) et il a fallu 4500 ans pour socialiser une méthodologie de l'interprétation (Hugues de St Victor) et en tirer profit (Descartes). Les efforts et inventions pour corriger les travers

de l'écriture sont définis par David Olson comme la *culture de l'écrit* (Olson, 1998). Cette culture s'avère déterminante : qui sait écrire sait penser, critiquer, inventer, et aussi imposer ses goûts et valeurs (Guichard, 2015).

L'internet renouvelle l'écriture, qui devient combinaison de 0 et de 1 (binaire) et en réseau. Ces transformations, ajoutées aux réflexions qui les ont générées (en partie liées à l'essor de la physique et de la mathématique) reconfigurent donc une nouvelle culture de l'écrit : la culture numérique. En bref, tout est à refaire : apprendre à manipuler de nouveaux systèmes de signes, à maîtriser les outils qui permettent de les happer par blocs et de les comparer, réinventer les systèmes de correspondances (culturels) entre l'écriture imprimée et l'informatique (bibliographies, références, etc.), réinventer des normes herméneutiques qui garantissent la compréhension par le plus grand nombre d'un même texte.

La chose n'est pas gagnée et nous vérifions aujourd'hui que de nombreux lettrés de l'ancien «régime» sont désarmés par le nouveau. D'autant que celui-ci est instable, du fait de la compétition des acteurs compétents pour ériger de nouvelles normes interprétatives. En matière de textes (formats de fichiers), de lois (droit des auteurs et des éditeurs), de sociabilité (réseaux sociaux), ou tout simplement d'«atomes scribaux» : j'entends par là les systèmes de signes minimaux comme les lettres de notre alphabet ou les signes écrits d'autres langues, qui nous permettent de composer des mots et des phrases — et qui sont eux-mêmes des combinaisons de 0 et de 1.

Là encore, les choix sont complexes car hérités de l'histoire et les résultats peuvent être surprenants.

3.3 Les atomes de l'écriture

Avant l'internet, les systèmes qui permettaient d'associer une lettre à une suite de 0 et de 1 (et donc à un nombre) n'étaient pas compatibles dès que l'on s'écartait d'une matrice minimale de 128 signes définie par une correspondance spécifique, nommée l'encodage `Ascii`. Un «é» écrit avec le système d'exploitation Windows ne donnait pas la même chose sur un Macintosh. Et Unix (l'ancêtre de Linux) développait une troisième norme. À ces normes s'en ajoutaient d'autres, pour les langues européennes autres que le français, pour l'arabe, et pour d'autres langues. Une nette amélioration apparut avec le web, pour les caractères accentués comme pour les sauts de ligne. L'internationalisation de l'informatique favorisa l'essor d'un projet utopique : encoder tous les caractères du monde, tous les signes de toutes les écritures : arabes, chinoises, cyrilliques, etc. Encore fallait-il s'entendre sur ce qu'est un signe ou une langue.

Le principe d'un encodage est assez simple : à chaque chiffre, il suffit d'associer un signe. Mais ce signe est-il un symbole, un caractère ? Sans entrer dans les détails, reconnaissons que cette question est complexe. Même en Europe, un caractère peut avoir plusieurs dessins, associés à ses polices. Une telle représentation s'appelle un glyphe. Par exemple un «T» écrit avec une police à empattement comme le Times diffère légèrement d'un «T» écrit avec une police sans empattement comme le Helvetica. L'encodage informatique est supposé ne pas en tenir compte : il se produit en amont, avant le travail du typographe. Par ailleurs, des symboles qui ne relèvent pas à proprement parler de l'écriture peuvent s'y insérer au fil du temps. Par exemple, le \$, qui n'est pas utilisé pour composer des mots, finit par devenir indispensable aux programmeurs (il définit souvent le statut d'une variable).

L'encodage actuellement privilégié s'appelle `UTF-8`. Il permet de reproduire facilement un signe en l'appelant par son «numéro». Les 128 premiers correspondent toujours à l'`Ascii` ; en revanche, les suivants divergent assez vite des anciennes conventions d'en-

codage, comme le `latin1`. L'UTF-8 permet d'écrire jusqu'à 2 millions de caractères⁹ même si on se sert surtout des 65000 premiers. Ce nouvel encodage est fort utile. Il permet aisément d'écrire en braille, berbère, grec, cyrillique, arabe, hébreu, perse, chinois, japonais, etc. et intègre la majorité des symboles mathématiques. Il produit aussi des résultats bizarres¹⁰, comme montré à l'URL <http://barthes.enssib.fr/art/testutf8.html>.

1. Il réintroduit la confusion caractère-glyphe qu'il est censé abolir. Par exemple, le dièse apparaît deux fois (codes 35 et 65119), l'esperluette (&) trois : codes 38, 64120, 65285. Des fractions, qui ne sont pas des signes mais des combinaisons de signes, apparaissent : comme le 1/3 (code 8531). On ne comprend pas trop pourquoi les chiffres romains, comme le VIII apparaissent comme un seul caractère (code 8551). De façon analogue, un «a» entre parenthèses (a) apparaît comme un seul caractère en position 9372. Les autres lettres de l'alphabet suivent. Il en est de même pour les nombres entre parenthèses ou suivis d'un point (9332 à 9371). Les symboles des centimètres et kilomètres ont eux aussi droit à leur encodage (13213 et 13214) alors que la juxtaposition du «c» et du «m» est essentielle pour retrouver les abréviations signifiantes.
2. La présence des ligatures arabes part d'une bonne idée : certaines lettres n'ont pas toujours la même forme suivant celles qui la précèdent ou la suivent. Pour autant, la majorité des logiciels qui permettent d'écrire en arabe savent gérer ces ligatures et la construction d'une police de caractères arabe suppose que chaque lettre a un dessin différent suivant sa position¹¹. Nous imaginons qu'un tel projet fasse écho à la tradition calligraphique arabe. Effectivement, apparaît une longue liste des ligatures triples, suivies de mots ou d'expressions condensés en un seul caractère : apparaissent alors le nom du dieu des musulmans (Allah, 4 lettres, code 65010), l'adjectif grand (akbar, 65011) ; Mohamed n'est pas en reste (65012). Nous découvrons que la première ligne du Coran (le *basmallah*) s'invite dans cet encodage (code 65021).

Ce point est problématique. Une fois que nous disposons d'un alphabet, jugeons-nous utile d'avoir un seul caractère pour écrire les trois mots «Liberté, égalité, fraternité», le «In God we trust» des Américains ou le «Je vous salue Marie» des catholiques ? Ici apparaît un étrange coup de force religieux dans un univers supposé faciliter l'écriture de tous les signes de toutes les langues. Et il n'aide pas à comprendre ce qu'est une écriture.

3. Enfin apparaissent des pictogrammes, qui n'ont aucune utilité dans la construction de nos phrases, sauf à construire des rébus : nous voyons proliférer des montres, sabliers, réveils, églises, cornues et parasols, jusqu'au logo d'Apple¹². Nous pouvons supposer que nos informaticiens avaient de l'humour. Mais là encore, ces «caractères» sont inutiles tant que des glyphes (des polices) n'ont pas été dessinés pour en faire un usage pertinent.

9. Dans la pratique, ce serait 4 milliards (2³²) ou 16 millions (2²⁴) suivant la manière de compter. Mais certains emplacement sont inutilisés et le début du code d'un caractère doit préciser combien d'octets ce caractère utilise. Nous visualisons la différence entre un encodage UTF-8 et un ancien (comme le `latin1`) quand nous voyons un «é» transformé en deux caractères (2 octets ancienne formule). Par exemple quand nous lisons «Institut Mines TÃ©lÃ©com» au lieu de «Institut Mines Télécom».

10. Une première étude des aberrations de l'UTF-8 a déjà été présentée (Guichard, 2016).

11. Pour une bonne introduction aux problèmes d'orientation et de ligatures de l'arabe sur le web, cf. <https://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Unicode/Test/arabe>.

12. Respectivement numéros 8986, 9203, 9200, 9962, 9879, 9969, 63743.

En conclusion, le système d'écriture produit est incompréhensible pour un linguiste ou un typographe et induit en erreur les néophytes en rompant avec la logique combinatoire propre à l'écriture. Il semble obéir à des logiques politiques où les personnes qui se sont invitées pour construire cet encodage ont négocié sa configuration en fonction de leurs intérêts et croyances. Certes, c'est ainsi que s'élabore dans un espace mondialisé toute norme technique. Mais nous ne pensions pas qu'il pourrait toucher les fondements de l'écriture, même si le gigantesque espace que se sont octroyés les fondateurs de cet encodage permettait d'y emmagasiner tous les signes qui leur passaient par la tête.

3.4 Le site Unicode

Nous pouvons nous offrir une idée des acteurs de cet encodage en consultant le site de leur consortium, intitulé «Unicode». Sur la page <http://www.unicode.org/consortium/members.html>, nous découvrons les principaux protagonistes de l'UTF-8 : Adobe, Apple, Facebook, Google, Huawei, IBM, Microsoft, Oracle, SAP et Yahoo (les multinationales de l'informatique et de l'internet), à qui s'ajoute... le ministère des Affaires Religieuses du sultanat d'Oman. Cf. figure 5. Nous comprenons mieux les intrusions de logos, de mentions religieuses ou l'introduction de pictogrammes naïfs dans un système qui se promettait de décliner de façon objective et désintéressée toutes les écritures du monde. Nous découvrons un système paré d'intentions généreuses mais incohérent, piloté par des organisations juges et parties qui sont de plus trop à l'écoute d'exigences religieuses. Au moins comprenons-nous la réalité du lien entre maîtrise de l'écriture et possibilité de façonner le monde suivant ses propres valeurs.

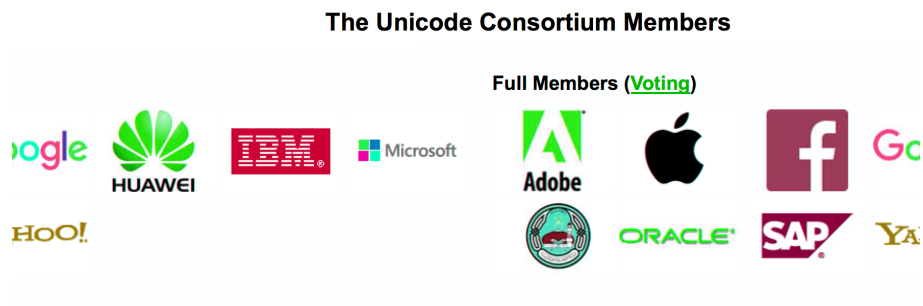


FIGURE 5 – La liste des membres avec droit de vote varie parfois légèrement au fil des mois. Le ministère des Affaires Religieuses d'Oman reste la seule institution étatique au côté des multinationales de l'écrit. Source : <http://www.unicode.org/consortium/members.html>, janvier 2016.

4 Conclusion

Nous réalisons que la possibilité de trouver des vérités ou de découvrir des preuves sur l'internet est complexifiée par les multiples opérations de brouillage rendues possibles par la jeunesse de cette technique. Nous en avons évoqué deux : les moteurs de recherche, qui détournent notre attention et nous mentent exagérément quant à leurs ressources. Le système d'encodage, qui intègre pêle-mêle caractères, pictogrammes, logos d'entreprises et divinités locales.

Nous pourrions donner d'autres exemples de la privatisation de l'internet et de l'écriture contemporaine, comme des surcharges cognitives et des capitulations politiques qui s'ensuivent (acceptation des surveillances publicitaire et policière). Notre intention n'est pas purement alarmiste : pour qui dispose d'une forte culture numérique et qui sait l'accorder avec une solide culture lettrée traditionnelle, l'internet est une technique intellectuelle, qui offre de nombreuses opportunités et qui donne de formidables idées. Des groupes très actifs et très compétents proposent des solutions cartographiques (OSM), informatiques (Linux), éditoriales (L^AT_EX), mathématiques et physiques (ArXiv, déjà cité) et des réflexions anthropologiques et philosophiques¹³ aussi intelligentes qu'altruistes.

Mais pour qui ne dispose pas d'un solide et double bagage intellectuel (classique et numérique, si possible en mathématique et en sciences sociales), ou qui se sent désemparé face à la nouvelle écriture, les choses sont d'autant plus difficiles que l'écriture contemporaine est d'une instabilité confondante, qu'aucun enseignement articulant technique et réflexivité n'est proposé (y compris à l'Université, sauf en quelques laboratoires méconnus comme le Pireh de Paris-I), et que l'industrie propose désormais des outils clés en main qui portent l'idéologie de leurs concepteurs : où les machines ont plus d'importance que les humains, où ces derniers sont au service des premières. Pour qui pensait qu'un automate pouvait alléger le travail de l'Homme, c'est un peu un comble.

Il n'est pourtant pas utile de désespérer. Si nous repensons notre rapport à la technique et les notions d'objectivité et si nous gardons en tête le fait que dans un monde illettré, les érudits imposent leur culture, nous pouvons aisément pratiquer et penser le monde «numérique». Une solution consiste à transmettre nos savoirs au plus grand nombre et à penser leurs histoires, leurs implicites idéologiques (Foucault, 1969), leurs effets. Et à écouter celles et ceux qui désirent partager leurs savoirs. L'idée d'une éducation populaire laïque et critique renaît ces temps-ci chez les détenteurs de la culture de l'écrit précédent ou contemporain.

13. Pierre-Antoine Chardel, Andrew Feenberg, les membres de l'équipe *Réseaux, Savoirs & Territoires*, etc.

Références

- CHAVALARIAS, D. (2012). La société (re)commandée. De la conjecture de von Foerster aux sciences sociales prédictives. In CHARDEL, P.-A., GOSSART, C. et REBER, B., éditeurs : *Conflits des interprétations dans la société de l'information : éthique et politique de l'environnement*. Lavoisier, Paris. URL : <http://halshs.archives-ouvertes.fr/hal-00632280/> ou <http://chavalarias.com/>.
- ECO, U. (1988). *Le signe*. Livre de poche-Labor. 1^{re} édition italienne : 1973.
- EDGERTON, D. (1998). De l'innovation aux usages. Dix thèses éclectiques sur l'histoire des techniques. *Annales Histoire, Sciences Sociales*, 4-5:815-837. Voir aussi *Des Sciences et des Techniques : un débat* (éd. de l'Ehess, 1998, pp. 259-287).
- FOUCAULT, M. (1969). *L'archéologie du savoir*. Gallimard, Paris.
- GOODY, J. R. (1994). *Entre l'oralité et l'écriture*. Presses Universitaires de France, Paris.
- GUICHARD, É. (1996). Internet, un coupable idéal? *Revue Terminal*, pages 181-190. n° 71-72, cf. <http://barthes.ens.fr/atelier/articles/guichard-nov-96.html>.
- GUICHARD, É. (2010). L'internet et l'écriture : du terrain à l'épistémologie. Habilitation à diriger des recherches. Université Lyon-1. URL : <http://barthes.ens.fr/articles/HDR-Guichard.html>.
- GUICHARD, É. (2014). Internet, technique et démocratie. In LEQUIN, Y. C. et LAMARD, P., éditeurs : *Éléments de démocratie technique*, Belfort. Presses de l'université de technologie de Belfort-Montbéliard.
- GUICHARD, É. (2015). Quelques apports de Jack Goody. les humains au prisme de leurs techniques. *Études Digitales*, 1. URL : <http://www.etudes-digitales.org>. Preprint : <http://barthes.enssib.fr/Goody/Hommage-Goody-EG.html>.
- GUICHARD, É. (2016). Écritures planaires : cartes, formules, codes et images. In AVENATI, O. et CHARDEL, P.-A., éditeurs : *Formes conçues, formes perçues. Pratiques du design et quête de sens en milieu technologique*.
- JEANNENEY, J.-N. (2005). Quand Google défie l'Europe. *Le Monde*, 22 janvier 2005.
- MATTHEWS, J. T. (2010). Quelques pistes de réflexion en vue d'une approche critique du web collaboratif. In MILLERAND, F., PROULX, S. et RUEFF, J., éditeurs : *Web social, mutation de la communication*, pages 329-340. Presses de l'Université du Québec.
- MOATTI, A. (2015). *Au pays de Numérix*. Presses Universitaires de France, Paris.
- OLSON, D. R. (1998). *L'univers de l'écrit*. Retz, Paris. Ed. orig. : *The World on Paper : The conceptual and cognitive implications of writing and reading* ; Cambridge University Press, 1994.