

Librologie 4 : Plain

Bonjour amis lecteurs et lectrices,

Un chapitre un peu exceptionnel pour la Librologie de cette semaine, dans de bien tristes circonstances. Je vous propose donc de réfléchir aujourd'hui au projet Gutenberg, à son fondateur Michael Hart... et à un objet mythique par excellence du monde geek : le *plain-texte*.

Bonne lecture à tous et à toutes !^[1]

Librologie 4 : Plain

Je précipite la publication de cette chronique (prévue pour le trimestre prochain) en apprenant à l'instant le décès de Michael Stern Hart, fondateur du Projet Gutenberg qui est la première (et sans doute la plus attachante) des bibliothèques Libres en ligne.



Au moment où j'écris ces lignes, le Projet Gutenberg s'apprête à fêter ses 40 ans – c'est-à-dire que sa naissance précède même celle du réseau Internet ! C'est le 1er décembre 1971 que tout commence, soit une bonne décennie avant l'apparition du mouvement Libre. Nous sommes à l'université d'Illinois, dont l'ordinateur central vient d'être mis en réseau avec une poignée d'autres, y compris (et c'est une grande première) au-delà du continent américain, pour former le réseau ARPANET.



Ce jour-là le jeune Michael Hart, âgé de 24 ans, va se retrouver devant ce joyau de technologie (un Xerox-SDS Sigma V, réparti dans quatre grandes armoires, doté de 64 Ko de mémoire et de 2 Mo de stockage sur bande), pleinement conscient de l'immense faveur qui lui est accordée (le temps d'ordinateur est précieusement minuté, la moindre minute ayant un coût exorbitant). Comment être à la hauteur de cet honneur historique ? C'est en chemin que l'idée lui vient, ayant récupéré à l'épicerie du coin un prospectus où est reproduite la Déclaration d'indépendance des États-Unis d'Amérique (lesquels États-Unis s'apprêtent alors à fêter leur bicentenaire) : il va frapper un grand coup, et partager avec le monde entier un patrimoine culturel qui lui survivra et traversera les siècles.

Le projet Gutenberg prend donc vie avec ce premier texte, qui sera alors consulté par un total de... six arpanautes. Cependant Hart voit loin (l'intitulé « Gutenberg » en témoigne), et se fixe pour objectif de mettre en ligne avant la fin du siècle, les 10 000 ouvrages les plus lus au monde. Si les débuts sont laborieux (isolé et privé de toute reconnaissance académique, Hart mettra vingt ans à atteindre seulement le chiffre de 100 livres), l'objectif sera non seulement atteint mais pulvérisé à partir des années 1990, avec l'avènement du Web et de la numérisation automatisée d'ouvrages imprimés. À l'heure où j'écris ces lignes, le projet Gutenberg approche les 40 000 opus (dont certains comptent plusieurs milliers de pages) ; cependant il a été rejoint par d'innombrables initiatives similaires ou concurrentes : de Wikisource à Google Books® en passant par Internet Archive, Gallica ou Européana, l'on ne compte plus aujourd'hui les bibliothèques en ligne dont le fonds est (plus ou moins) librement accessible au public. Le réseau Internet tout entier s'est massivement transformé en vecteur de diffusion et de consommation culturelle ; P2P, MP3,

streaming, podcasts, multimédia en tout genre, culte de l'image (particulièrement de l'image animée). Les livres eux-même se sont faits *e-books*, terme dont j'attends encore que l'on m'explique ce qu'il peut bien signifier et pourquoi on a cru bon de l'inventer – si ce n'est pour vendre, à l'occasion, des attrape-gogos numériques.

Naturellement, le projet Gutenberg n'a pas été insensible à cette évolution (qu'il a largement contribué à susciter). L'on trouvera ainsi sur la page d'un livre au hasard, la mention « *eBook* » largement mise en avant, ainsi que les liens suivants : Bibrec (métadonnées bibliographiques), QR Code (« Flashcode » bi-dimensionnel pour téléphone mobile), Facebook, Twitter (sans commentaire), HTML, EPUB, Kindle, Plucker, Qi00 Mobile... L'on trouvera aussi, à l'occasion, des versions audio des livres en question, dans des formats Libres (Ogg Vorbis, Speex) et non-libres (MP3, Apple iTunes). On le voit, la priorité du projet Gutenberg est d'être visible et accessible (commodément) au plus grand nombre, même au prix de quelques « fautes de goût » Libristes.

Hart assume pleinement cette diversité :

Le projet Gutenberg est animé par des idées, des idéaux et un idéalisme.

Le projet Gutenberg n'est pas animé par un pouvoir financier ou politique.

De ce fait le projet Gutenberg est entièrement animé par des bénévoles.

Étant entièrement animés par des bénévoles, nous sommes réticents à toute forme d'autorité sur nos bénévoles, pour leur dire ce qu'ils doivent faire ou comment ils doivent le faire.

Nous offrons autant de libertés que possible à nos bénévoles, dans le choix des livres à partager, des formats dans lesquels les partager, ou toute autre idée qu'ils pourraient avoir quant à « la création et diffusion d'eBooks ».

Le projet Gutenberg n'a que faire d'établir des standards. Si c'était notre rôle, nous aurions accepté avec joie la proposition qui nous a été faite de convertir nos eBooks en HTML lorsque le Web était une idée toute nouvelle en 1993 ; nous nous satisfaisons d'apporter des eBooks à nos lecteurs dans tous les formats que nos bénévoles souhaitent produire.

(...) Nous encourageons les gens à nous faire parvenir des eBooks dans n'importe quel format, puis nous cherchons des bénévoles pour les convertir dans d'autres formats, et corriger peu à peu les erreurs d'édition.

(...) Nous voulons présenter au monde autant d'eBooks, dans autant de formats et en autant de langues, que possible.

Quelques formats exotiques ou propriétaires qu'il puisse proposer, rien ne me semble mieux illustrer la démarche d'accessibilité et d'universalité du projet Gutenberg que son choix, encore réaffirmé aujourd'hui, de proposer tous ses ouvrages sous forme de simples fichiers texte : dépourvus de mise en forme, d'illustrations ou de toute fioriture, ces fichiers symbolisent toute une mythologie Libriste remontant aux débuts de l'informatique, et qui est celle du *plain-texte*.

L'expression « plain text » est de celles que l'on comprend aisément sans toutefois parvenir à la traduire de façon entièrement satisfaisante. Le projet Gutenberg propose « *texte brut* » ; Mozilla Thunderbird, sous la plume de Cédric Corazza, y ajoute « *texte normal* », plusieurs éditeurs en ligne proposent « *texte pur* », l'interface de Google Mail® indique « *texte seul* ». Wikipédia est tout aussi désemparé, juxtaposant allègrement, inspirez profondément, « *fichier texte ou fichier texte brut ou fichier texte simple ou fichier ASCII* »... et encore, ce n'est qu'après



avoir rejeté la traduction littérale « plein texte », qui n'était pourtant pas la pire : le mot anglais *plain* (lui-même dérivé du vieux français) se situe en quelque sorte à mi-chemin de nos adjectifs « plan » et « plein ». Aucune de ces traductions, hélas, ne rend l'idée de dépouillement, de linéarité et (par association d'idées) de plénitude exprimée par le vieux mot français **plain**, qui me semblerait pourtant avoir là une magnifique occasion d'être remis en usage : ne pourrait-on pas parler de plain-texte comme l'on parle de plain-chant ?

Le plain-texte est, donc, le premier apport historique de l'informatique : pendant plusieurs décennies, les ordinateurs n'ont permis de n'échanger que cela. C'est aussi le plus fondamental – comme l'avait bien compris Michael Hart – et le plus irremplaçable : encore aujourd'hui, l'essentiel des communications numériques entre humains se fait sous forme textuelle. Courrier électronique, messagerie instantanée, commentaires sur le Web... Le plain-texte connaît même, ces dernières années, un surprenant regain d'intérêt sous une forme minimale, avec la mode des micro-blogs dont nous serons amenés à reparler ici.

Il est également l'élément immatériel le plus versatile et le plus plastique : il peut servir à écrire des textes, des livres, des programmes ou des fichiers de configuration, à éditer des documents scientifiques ou des partitions musicales, voire à décrire des graphismes, des objets en trois dimensions, ou encore... à peu près tout ce que l'on veut.

Le plus robuste aussi, sans nul doute : un fichier texte corrompu ou tronqué a de meilleures chances d'être reconstitué qu'un fichier binaire. À l'utilisateur humain, le plain-texte demeure accessible et intelligible ; rassurant, somme toute. Le programmeur Douglas McIlroy, parrain de la philosophie Unix il y a un demi-siècle, ne disait pas autre chose : « *n'écrivez que des programmes dont le format d'entrée est du texte pur, car c'est là une interface universelle.* » Universalité certes

très relative à l'époque, puisque le codage utilisé ne permet alors l'utilisation que d'un nombre restreint de caractères ; il faudra attendre les décennies suivantes pour conquérir le bas-de-casse (comme nous l'évoquons dans notre chronique sur rms), puis les caractères accentués et l'Unicode (d'ailleurs encore chaotique aujourd'hui)...



Le plain-texte est, enfin, l'âme de l'informatique : binaire et code machine étant réservés aux tréfonds du bas-niveau des calculateurs, c'est sous forme de langages de programmation textuels, toujours plus naturels, que les programmeurs s'adressent et commandent à l'ordinateur... et même les simples

utilisateurs qui, aujourd'hui encore, ont la patience d'apprendre à se servir d'interface textuelles, voient leur approche de l'informatique changée à jamais : dialoguer avec l'ordinateur en ligne de commande, c'est comprendre sa logique ; c'est intervenir directement dans son fonctionnement, sans la médiation factice et opaque d'une interface conçue par des humains pour un utilisateur déshumanisé, théorique et paresseux.

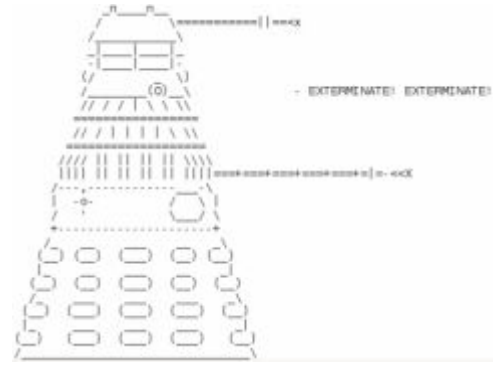
En un mot, le plain-texte est *Libérateur* : de fait, le mouvement du logiciel Libre ne milite pas pour autre chose, depuis trente ans, que d'avoir le droit d'accéder au code source, c'est-à-dire aux programmes sous une forme de plain-texte lisible et intelligible, modifiable et aisément partageable. La parenté avec la démarche du projet Gutenberg n'en est que plus frappante.

Libérateur, accessible,... mais également respectueux : un document en plain-texte ne s'impose pas à son lecteur, n'exige pas telle ou telle manière d'être lu. Il permet (dans un environnement graphique ou même en mode console) de choisir sa propre taille de texte, sa propre police (l'on préférera en

général une fonte à espacement fixe, comme pour une machine à écrire). À l'époque où se développent les premières interactions sociales en ligne (*e-mail* dans les années 1970, *Usenet* dans les années 1980), la question ne se pose pas en ces termes : les limitations techniques (protocoles limités à l'ASCII, lenteur et coût des communications) imposent d'aller au plus court. C'est l'hégémonie, dans les années 1990, non seulement de l'HTML mais de l'informatique grand-public dite « personnelle » et des interfaces propriétaires qui remettra en cause, ô combien, ce modèle : advient le règne de la vulgarité, dont le parangon sera le mail en HTML, abomination que nous subissons encore aujourd'hui.

L'attitude, intègre et exigeante, qui consiste à préférer le plain-texte aux merveilles *bling* du texte soi-disant « riche », est souvent qualifiée d'élitiste ou de puriste (comme en témoigne l'expression « texte pur ») par ses détracteurs. Il ne s'agit pourtant aucunement de remettre en cause l'appropriation des outils informatiques par le plus grand nombre, ce dont on ne peut que se réjouir. Il s'agit d'une simple question de culture : de même que l'on attend de tout citoyen qu'il possède quelques références culturelles (à commencer par un minimum de grammaire écrite, ou une connaissance de base du code vestimentaire occidental), il est légitime de rappeler que notre culture d'internautes contemporains (faite de *smileys*, de *memes*, de *flood*, de *flamewars* et de *troll*) est le fruit de plusieurs générations de geeks qui l'ont construite sous forme purement textuelle. (Je m'empresse de préciser ici que je n'appartiens pas moi-même à cette génération, n'ayant pas eu l'usage d'un ordinateur avant le XXI^e siècle.)

Rappeler également, car on l'oublie trop souvent, la puissance expressive du plain-texte. Je ne reviendrai pas ici sur *l'ASCII Art*, même si c'est un exemple frappant. Je suis davantage intéressé par la minutie avec laquelle les codeurs (programmeurs ou non) rédigent leur code source, développant



de véritables traditions (*coding styles*) à la fois pratiques et esthétiques dans lesquelles il me semble voir une forme de langage expressif, métalinguistique et poétique – sur laquelle je ne m'attarderai pas davantage ici. Comme toute forme de communication, le plain-texte a subi au fil des décennies une exigence d'expressivité, et certaines traces en sont particulièrement visibles d'un point de vue formel : écrire en capitales signifie que l'on crie, insérer un smiley en fin de phrase dénote un ton ironique, et ainsi de suite.

Cette exigence d'expressivité influe même sur la structure du langage, qui se surcharge de signes. Un bon exemple en est la longueur des phrases, qui à l'oral importe bien moins que le ton et le propos : dans un message électronique en plain-texte au contraire, rien n'est plus cassant qu'une phrase courte et lapidaire, et l'on se surprendra fréquemment à rallonger artificiellement ses phrases pour ne point froisser son interlocuteur. Les effets de juxtaposition, également, sont frappants – particulièrement dans un courriel auquel l'on répond de façon entrelacée avec le message d'origine. Autre effet de juxtaposition – dont j'use abondamment dans ces chroniques sous forme d'hyperliens –, l'insertion d'URLs dans le discours. Enfin les retours à la ligne et effets typographiques, pour rudimentaires qu'ils soient, permettent une certaine dramatisation (au sens de dramaturgie) du discours ; je pourrais en donner un exemple.

Ici même.

Mais à quoi bon ?

hausse les épaules

Enfin bref.

Il ne serait donc pas totalement honnête de théoriser, comme je l'avais moi-même fait dans une première version de cet article, que le plain-texte permet d'ignorer la forme pour se concentrer sur le contenu. La forme est toujours présente, si discrète soit-elle ; la différence est qu'elle ne fait pas nécessairement sens par elle-même, et demande un minimum d'attention, autant à l'émetteur qu'au récepteur, pour lire, littéralement, *entre les lignes* de plain-texte.



L'illustration idéale de ce propos m'est, à nouveau, fournie par Michael Hart, virtuose du plain-texte s'il en fut – et Libriste engagé. Sur sa très modeste page web, il évoque quelques-uns des (trop prévisibles) ennuis juridiques qu'a pu subir le projet Gutenberg dans la dernière décennie, certains éditeurs traditionnels s'étant manifestement lancés dans une croisade contre le domaine public et l'intérêt général. Il faut lire les réponses rédigées – toutes en plain-texte, évidemment – par Hart et son bras droit Gregory Newby aux mises en demeure d'éditeurs outragés : à la fois polies, implacables... et d'une ironie mordante : un exemple à suivre. La réponse de Hart concernant le livre *Anthem*, en particulier, est à lire à tout prix : outre son flegme et son exactitude juridique, chaque ligne fait exactement le même nombre de caractères, le tour de force par excellence de tout geek qui se respecte...

Telle est la leçon, et le souvenir, que nous laisse Michael Hart : celui d'une époque, d'un esprit où la rigueur se mêle à l'ambition et la fantaisie, où l'économie de moyens n'empêche point l'élégance et l'humour, et où, enfin, l'attrait de la nouveauté ne laisse pas s'estomper l'Histoire et le genre

humain dans son ensemble.

To Michael S. Hart, a plain human being.

Notes

[1] Crédit photo : 1. Michael Hart 2. Xerox-SDS Sigma 5 – Laughing Squid – CC by-nc, 3. Plain text – oxygen-icons.org – LGPL, 4. It ain't what it used to be – revdode – CC by-nc-sa, 5. ASCII Art et 6. Michael Hart et Gregory Newby – Marcello – GFDL.