



Données et traces numériques (sous rature)

Catégorie : **Mundus Numericus**

Tags : **art, écologie, futur, information, infrastructures, philosophie, physique, société, technique**

Personnages : **Mohamed Atalla, Jacques Derrida, Dawon Kahng, Jean-Pierre Nadau, Herman Parret, Trevor Paglen**

18 janvier 2020

Les termes « données », « traces », etc. se sont glissés comme des œuvres de faussaires dans la novlangue philosophico-numérique. Analysons et démasquons !

Au détour d'un sentier de brousse, nous avons croisé cette information : « *Le volume de données mondial sera multiplié par 45 entre 2020 et 2035* »¹. Ainsi, partant de 50 zettaoctets (ZB) de données en 2020 (1 ZB = mille milliards de milliards d'octets), on atteindrait 175 ZB en 2025 et plus de 2000 ZB en 2035, soit une multiplication par 3,5 tous les cinq ans. Cette croissance, tirée principalement par l'émergence du milieu naturel digital, est peu surprenante. Mais qu'en est-il du volume proprement dit ? Est-ce « beaucoup » ? Nous proposons de le re-présenter autrement de façon à pouvoir... l'observer ! Sous cet angle, précisément, que voit-on ? Précisons pour la suite que les termes « donnée », « trace » et « information » seront à peu près interchangeables et devront être compris dans leur sens habituel.

¹ Marjolaine Tasset pour journaldunet.com - 17 mai 2019 - [Le volume de données mondial sera multiplié par 45 entre 2020 et 2035](#)

Le long des mémoires

Le langage est toujours en retard d'une technique mais Mundus Numericus a émergé si vite que le décalage est inédit. De quoi parle-t-on en évoquant le « *volume de données mondial* » ? Une *image* nous vient-elle ? Il faudrait procéder par associations d'idées, par métaphores, mais rester alors prisonniers d'une langue plus ou moins adéquate. Alors pourquoi ne pas regagner l'inframonde physique dont les concepts sont supposés rendre compte afin de le re-présenter différemment ? L'artiste Trevor Paglen avait suggéré ce genre de mouvement en nous proposant de *regarder* le trafic internet en tenue de plongée [le long des câbles sous-marins](#) (nous avons au passage ressenti l'immense fragilité des constructions numériques masquée derrière l'habituelle narration de la puissance et l'omniprésence du « virtuel »).



Trevor Paglen le long des câbles sous-marins

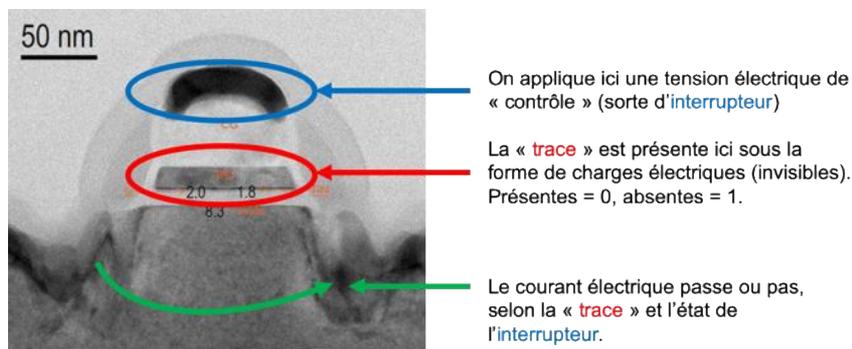
Où sont donc nos « données » ? En quoi *physiquement* consistent-elles ? (Tâchons ici de garder le lecteur non technicien, auquel nous tenons beaucoup, sans choquer le spécialiste que nous tenterons de rattraper par quelques notes de bas de page ...). On suppose donc que l'engramme de l'ensemble mondial des données et traces numériques est inscrit exclusivement dans une vaste mémoire de type « flash » (celle qui équipe nos smartphones ou nos clés USB par exemple)². Dans cette mémoire sont inscrits les atomes d'information, les « bits » (0 ou 1), qui peuvent être lus et effacés. Il s'agit donc d'une sorte d'immense « ardoise » sur laquelle nos artefacts numériques font des « dessins » et laissent des traces. Rappelons déjà cette évidence : *ce n'est pas l'humain qui laisse des traces numériques*. « Nos traces », « nos données », sont évidemment des vues de l'esprit c'est-à-dire des expressions du langage.

² Il existe encore de nombreux équipements dotés de supports classiques (disques durs etc.). Mais nous considérons que l'approximation proposée (tout est en mémoire NAND par exemple) est assez acceptable.

Un objet extraordinaire

L'atome de cette mémoire est un objet extraordinaire : le transistor MOSFET (« *Metal-Oxide-Semiconductor Field-Effect* »). Il a été conçu en 1959 aux laboratoires Bell par un ingénieur égyptien, Mohamed Atalla, et un ingénieur coréen, Dawon Kahng. Ils n'ont pas reçu de prix Nobel pour leur invention mais sont régulièrement acclamés pour cette contribution essentielle à notre milieu numérique. Le transistor MOSFET pourrait d'ailleurs figurer au Guinness Book des records au titre d'objet le plus fabriqué de tous les temps : 13 000 milliards de milliards ($1,3 \times 10^{22}$) entre 1964 et 2018, soit 13 trilliards ou 13 « zettas ». L'ordre de grandeur du « *volume de données mondial* » est bien confirmé.

Un type particulier de transistor MOSFET, dit « à grille flottante », équipe les mémoires « flash ». Voici à quoi il ressemble et où sont donc en réalité les traces numériques³ :



Transistor MOSFET à grille flottante

En haut à gauche figure l'échelle de l'objet : 50 nanomètres, 50 milliardièmes de mètre, proche de l'échelle atomique (dixième de nanomètre). Le transistor MOSFET est extrêmement petit et les industriels⁴ savent les graver côte à côte par milliards⁵. En supposant que chaque transistor tienne sur un carré de 200 nanomètres de côté (densité moyenne des mémoires fabriquées en 2019⁶) et contienne un bit d'information⁷, le volume mondial de données tient sur un carré de 100 km de côté. Voilà à peu près la taille de l'ardoise sur laquelle sont inscrits déchets et dessins numériques, les 50 ZB de données mondiales : la surface du département de la Gironde couverte de transistors MOSFET.



Supposons pour finir que cette surface ait la même résolution qu'un écran de smartphone, dont les pixels font environ 0,1 mm de côté et « portent » 24 bits d'information (les couleurs). Chaque pixel du

³ Phys.org – 11 décembre 2012 – [Ultra-thin hybrid floating gate cell presented at IEDM2012](#)

⁴ Comme les coréen Samsung et SK Hynix, les américains Intel et Micron, le japonais Toshiba... Il y a aussi des chinois. Aucun européen à l'horizon...

⁵ Et maintenant les uns par-dessus les autres, en « 3D », avec plusieurs dizaines voire centaines de couches superposées.

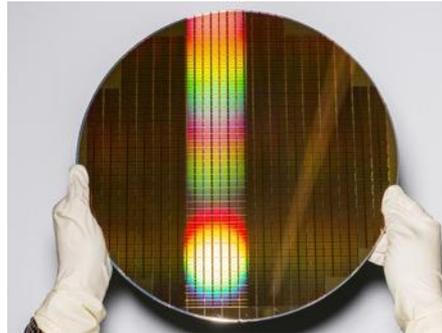
⁶ Wikipédia – [Transistor count](#)

⁷ Un transistor MOSFET peut aujourd'hui stocker plus d'1 bit d'information : 2, 3 ou 4, peut-être davantage à l'avenir.

département, couvert de 250 000 transistors MOSFET, est donc 100 fois mieux résolu que l'écran d'un smartphone.

La couleur de l'ardoise

Malheureusement, nous ne pouvons pas voir cette ardoise « à la Trevor Paglen ». Elle est découpée en milliards de morceaux répartis dans nos ordinateurs, smartphones, objets... ainsi que dans les équipements des 4000 datacenters mondiaux. Mais faisons-nous quand même le plaisir d'imaginer le département de la Gironde recouvert par cette surface irisée :



Un « wafer » de mémoires à transistor MOSFET

Les traces et données sont ainsi *en dernier lieu* des signatures électroniques nanométriques sur ce substrat. Tout concept et discours au sujet de l'information mondiale, tout langage au sujet du numérique, doit être rappelé à cette matérialité-là. C'est la puissance propre d'une re-présentation artistique ou imaginaire : elle est infra-conceptuelle et nous permet de bousculer une langue qui devient inadéquate (« donnée », « trace » ...). Mais continuons l'exploration.

Avec un septième sens

Si nous disposions d'un sens nous permettant de « voir » les dépôts électroniques des transistors MOSFET et donc les bits d'information, à quoi ressemblerait un petit morceau de cette surface ? Peut-être à une œuvre de l'artiste Jean-Pierre Nadau :

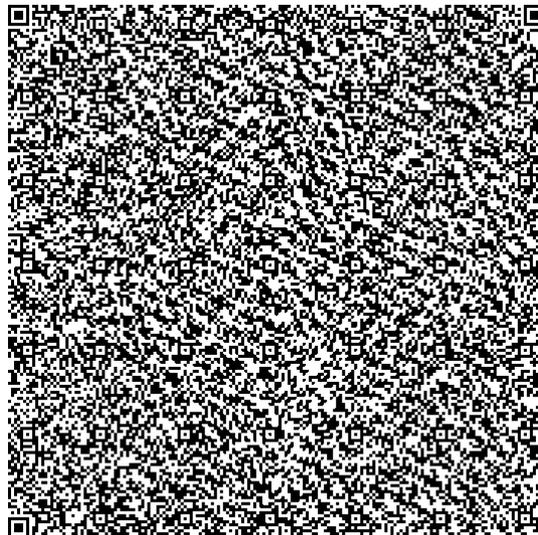


Dessin à l'encre de chine sur papier, 21x30 cm

Non pas ! Même vu de loin, le blanc et le noir juxtaposés de façon apparemment anarchique dessinent de vagues lignes, un vague centre, déclenchent de furtives sensations et interprétations. Il n'y a pas d'ordre apparent et malgré tout il ne s'agit pas d'un désordre total. Nous y voyons des déchets, des traces inconscientes et automatiques, un travail au sujet duquel Jean-Paul Gavard-Perret écrits :

L'amoncellement des détails crée des permanentes au chaos pour d'une certaine manière et comme disait Zazie dans le métro, « le rendre indéfrisable ».

Le rendre « *indéfrisable* », c'est-à-dire lui donner une consistance et une résistance. Mais nous voyons bien que ces « permanentes au chaos », sur notre ardoise numérique, ne sont permanentes que pour l' « artiste » qui signe les charges électriques des transistors MOSFET. Cet « artiste », c'est l'ensemble mondial des équipements numériques, et il ne comprend que lui-même. 1 micron carré de son « œuvre » ressemble plutôt à ceci, de la neige de téléviseur qui s'anime par saccades :



Il n'y a rigoureusement aucun moyen d'exploiter ces traces (ni même d'observer la moindre régularité) sans ce monstrueux appareillage : le système numérique mondial et ses algorithmes. Cette mémoire est littéralement *inhumaine*. De là, survolant cette Gironde irisée, nous pouvons tirer au moins cinq observations pour des réflexions à venir.

1 - une mémoire sans « valeur »

Encore un peu d'imagination. Faisons appel au bien serviable extra-terrestre, qui découvre cette ardoise (inerte) et accède par un moyen quelconque, comme ce septième sens, aux transistors MOSFET et à leurs dépôts électroniques. Il observe ainsi qu'il existe deux sortes de dépôts (0 ou 1) et tente donc de découvrir un ordre, une structure associée à ces deux espèces. Il doit « lire » les 50 ZB de données actuelles, soit 4×10^{23} transistors. E.T. remarque au passage que 4×10^{23} est un nombre similaire à une « mole », la quantité d'atomes dans 27 grammes d'aluminium. Au fond pas grand-chose, se dit-il...

Première remarque : s'il n'arrive pas à découvrir la moindre structure (« à la Nadau ») sur cette surface moirée, il sera face à un objet dont la thermodynamique nous indique qu'il ne tirera même pas un wattheure⁹, tout juste de quoi allumer une ampoule de 20 watts pendant 1 minute (un peu décevant pour l'ensemble mondial des « informations »...).

Seconde remarque : notre extra-terrestre souhaite « lire » chaque transistor afin de découvrir le sens de tout ceci, s'il y en a un. Muni de techniques de lecture similaires aux nôtres, il doit mettre en marche son « groupe électrogène » et consommer (comme nous) 1 microjoule par transistor lu, soit 31 GWh au total pour lire toute l'ardoise, l'équivalent de la production électrique de 31 réacteurs nucléaires pendant 1 heure. C'est ainsi l'énergie nécessaire pour lire *une fois* toute l'information mondiale¹⁰. Le jeu a intérêt à valoir la chandelle pour E.T. car une fois lus et relus, ces « bits » devront être corrélés par des algorithmes, puis les résultats mis en mémoire ailleurs etc. et la consommation d'énergie explosera.

Les « griffures » électroniques sur cette ardoise de 10 000 km² ne sont pas comme les lettres imprimées des ouvrages d'une bibliothèque, les empreintes de pas sur un sentier, les traces ADN sur une scène de crime... Il n'y a matériellement aucun moyen de les exploiter, maintenant ou plus tard, sans les innombrables couches d'algorithmes, de réseaux, de cartes électroniques, d'appliances. Aucun E.T. ne pourra *matériellement* nous aider à reconstituer la moindre information si notre machine à interpréter défaille. Tout juste pourrait-il en tirer de quoi allumer brièvement une ampoule et puis plus rien.

2 – Faut-il faire le « procès » des théories de l'information ?

Cette surface, comme un palimpseste, finit par s'altérer à force de lectures et écritures. Les transistors se détériorent. Il faut donc l'énergie d'une sorte de Sisyphe pour la rénover en permanence, recopier et dupliquer les « données » à coups de GWh.

L'espèce humaine est extrêmement astucieuse dans son maniement de la matière (ici le transistor MOSFET) et en cela réside son absolu génie : le « bricolage ». Or, cet ingénieur bricolage, qui anime les principes des théories de l'information, produit quantité de déchets sur une ardoise illisible sans « centrale à charbon ». D'où ce doute profond pour un professionnel des systèmes d'information : les théories de l'information elles-mêmes, dont on fait aujourd'hui l'exégèse en admirateurs inconditionnels, doivent-elles être rejetées comme inefficaces, voire dangereuses ? Les « 0 » et les « 1 », que plus personne n'interroge (même les « quanticiens »), sont-ils complètement idiots ?

Nous n'avons rien de particulier à proposer mais donnons l'ordre d'idée de notre désarroi : nous pourrions rêver, par exemple, d'une « bio-information » dont la représentation et la manipulation (écriture, lecture...) relèverait directement de cycles éco-naturels dans un biotope.

⁹ Le raisonnement fictionnel consiste à considérer l'entropie S d'un système à 2 puissance 4×10^{23} degrés de liberté en appliquant la formule de Boltzmann, ce qui donne $S = 3,8 \text{ J/K}$ soit 1 100 Joules = 0,3 Wh à la température ambiante.

¹⁰ C'est 10 fois plus pour tout réécrire et 100 fois plus pour effacer l'ardoise !

3 – L’algorithme comme « thérapie »

Nous stockons une mémoire sans valeur, donc nous ne souffrons pas d’un excès de données et de traces, loin de là, mais d’un *excès d’algorithmes* ; non pas d’un excès de mémoire mais d’une compulsion de *remémoration*. C’est en grande partie culturel. L’oubli est une faute morale et nous devons passer notre temps à nous re-présenter notre mémoire. Le philosophe Herman Parret écrivait ceci¹¹ :

Il est vrai, nos philosophies ne sont pas favorables à l’oubli. Freud : la psychanalyse sert à remplir les trous de la mémoire et constitue ainsi sa thérapie.

Nous venons de voir, grâce à E.T., que la mémoire numérique totale a une valeur (future) intrinsèque très faible. C’est l’algorithme qui, animé par une compulsion de re-présentation, crée de la valeur en « remplissant les trous » de la mémoire individuelle et collective : *l’algorithme est, si l’on ose dire, une thérapie*. Le problème est que le thérapeute est le système du pouvoir (technique, économique, financier voire politique) et n’a pas nécessairement intérêt à notre « guérison ».

4 – Les dégâts de l’hypermnésie

Le terme d’ « *hypermnésie* » est souvent utilisé pour qualifier la mémoire numérique, en particulier celle du web. Cette image est finalement assez juste à condition de la comprendre non pas comme métaphore d’un excès de mémoire (sans valeur) mais bien d’une authentique hypermnésie clinique¹² : « *temps excessif consacré à se remémorer son passé* », exécution permanente d’ « *associations incontrôlables* » (l’IA et le big data), etc. Notons cependant que l’hypermnésique n’a pas de capacité de mémorisation supérieure à la moyenne¹³. Il souffre bien d’une pathologie de la remémoration.

Cette pathologie algorithmique, entretenue par 10 000 km² de mémoire flash en constante augmentation, semble provoquer au moins deux sortes de dégâts collatéraux. Tout d’abord, les « fake news » (la valeur de vérité d’une information dépend plus que jamais du contexte algorithmique de sa re-présentation – or, la re-présentation *ne cesse jamais* et érode ainsi la vérité) et ensuite, plus insidieusement, les dérives de ce que nous appellerons provisoirement le « re-jugement », c’est-à-dire le regard moral porté *aujourd’hui* sur la re-présentation au pixel près de faits anciens gravés électroniquement dans les transistors MOSFET. Ce second phénomène est fascinant. Le monde numérique est un tribunal permanent où n’existe ni le salut *civilisé* de la prescription (la mémoire est ineffaçable et le droit à l’oubli un simple droit à la « non-lecture »), ni la médiation *civilisante* d’un récit historique. Il s’agit de cet état que l’investisseur [Ray Dalio](#) avait qualifié de « transparence radicale ».

Allons-y pour un clin d’œil à l’actualité. Donc « *ok boomer* » par exemple. Ok ? Mais chaque génération sera accusée par ses enfants, par cette double « malédiction » de la non-prescription et de la non-médiation humaine par l’histoire : ce qui est mal aujourd’hui fut

¹¹ Herman Parret in Protée, 32 (2), 37-46 – 2004 – [Vestige, archive et trace : Présences du temps passé](#)

¹² Wikipédia – [Hypermnésie](#)

¹³ Contrairement au monde numérique ; ici réside la limite de la métaphore.

donc de tous temps une faute. L'ardoise et la re-représentation algorithmique de faits intacts, garantis MOSFET, dans tout leur éclat haute définition, nous font oublier les vertus du temps qui passe et de ces inventions bien humaines que sont l'histoire et la prescription. Faudra-t-il donc trouver d'autres procédés civilisants pour résister aux effets déchirants de l'hypermnésie ?

Tentons une boutade affectueuse, un « *ok number* » ou n'importe quoi d'autre qui signifiera « *shut up* », pourra être renvoyé par leurs propres enfants à la génération d'aujourd'hui qui aura a) éteint les libertés et b) ajouté sa contribution au dérèglement climatique. Il suffira de rappeler dans 20 ans des zettas de traces MOSFET et de les présenter au « tribunal » :



C'est évidemment ridicule et très injuste. Nous sommes simplement tous humains, trop humains. Conclusion : la solidarité intergénérationnelle devient un nouvel impératif ; il faudra probablement renforcer notre capacité à « pardonner » puisque, semble-t-il, le « thérapeute » a décidé que nous n'oublierions pas, qu'il fallait boucher tous les trous de l'histoire et nous la re-présenter sous la forme d'un éternel présent hyper-précis.

5 - « Données » et « Traces » ? Oui, mais sous rature...

Ce ne sont pas les humains qui laissent des traces sur l'ardoise, sinon nous les reconnâtrions. Ces traces sont des *dessins* réalisés par des machines que nous ne pouvons pas comprendre et à qui nous faisons nécessairement *confiance* pour assurer l'intégrité entre ce que nous avons voulu signifier et ce qui sera restitué. Il n'y a, si l'on peut dire, aucune « trace de donnée », au sens où nous entendons « donnée », sur l'ardoise. *Une donnée est une fiction* mais nous faisons comme si c'était une réalité accessible à une proposition « ontologique ». Ainsi, une information « circule », des traces sont « laissées » ou « captées », une donnée doit être « protégée » ou encore « propriété de », etc. La « donnée » au sens large (information, trace...) a ainsi une existence propre attestée par le langage et montrée par les traitements algorithmiques, le plus immédiat étant l'affichage à l'écran. Or, une donnée n'est pas une chose mais une pure convention de langage.

Le philosophe Jacques Derrida a développé l'usage de cet « outil » philosophique et typographique génial inventé par Martin Heidegger : la mise « *sous rature* ». Ce procédé, qui a inspiré de nombreux artistes¹⁴, consiste à barrer un mot tout en laissant à sa place dans le texte, comme ceci : une donnée est une fiction. Cette convention signifie que le mot sous rature est « *inadéquat bien que nécessaire* » ; il n'est pas adapté au concept qu'il représente mais doit continuer à être utilisé tant que le langage ne propose rien de mieux. Nous devrions ainsi écrire : l'information circule, des traces numériques sont laissées, la propriété des données est une illusion, etc. En effet, à quelques nuances près, aucun de ces termes « anciens » n'est connoté par le fait qu'il s'agit désormais de *fictions chaque fois reconstruites par exécutions d'algorithmes*, de résultats calculés à partir d'une « *mémoire sans valeur* ». Peut-être certains auteurs ont-ils déjà les mots adéquats qui éviteraient d'énoncer sans gêne apparente ces locutions inconfortables (« protection des données », « propriété des données »...) qui elles, en revanche, n'ont malheureusement rien de fictions juridiques.

En avant !

Voyons-nous déjà mieux l'inadéquation radicale de cette expression : « *volume de données mondial* » ? Toujours est-il que ce volume va donc « *exploser* » en grande partie à cause des objets connectés (les tuyaux 5G arrivent). Si l'on en croit les projections du cabinet IDC, il sera multiplié par 45 d'ici à 2035 (autant de bits que d'atomes dans 1 kg d'aluminium). L'« ardoise » (dans tous les sens du terme...) couvrira alors à peu près la surface de la France. Mais c'est à la flambée algorithmique qu'il faudra faire face, qui seule peut donner de la valeur à ce « kilogramme » de matière en calculant sans cesse données, traces et informations. L'enjeu énergétique est considérable.

Cette flambée convoque au minimum ces réflexions tous azimuts : la mise en cause des théories actuelles de l'information, la solidarité entre les générations (puisque le temps ne passe plus), l'assistance de la philosophie, qui ferait œuvre utile en sous-raturant un langage auquel nous nous sommes habitués mais devenu inadéquat voire trompeur.

¹⁴ Voir à ce sujet le texte de Raphael Rubinstein – [Under erasure](#)

Post-scriptum « thermodynamique »

Le système numérique mondial peut être envisagé comme une machine « thermodynamique » qui produit du « travail » sous la forme de données, traces, informations fournies à la population mondiale des humains mais aussi, désormais et davantage encore, aux artefacts : objets connectés, objets autonomes, « intelligences » artificielles, villes « intelligentes » etc.

La transformation des traces électroniques des transistors MOSFET en « travail » par les algorithmes consomme autour de 1000 TWh d'électricité (consommation électrique de l'Inde) et représente 3% à 4% des rejets mondiaux de tonnes équivalent CO₂ ([L'illusion d'un monde numérique vert](#)). Les données, traces et informations sont loin d'avoir une existence propre et leur « fictionnalisation » dans le grand théâtre de nos smartphones, tablettes, montres, etc. coûte cher... Espérons malgré tout que les avantages l'emportent sur ces inconvénients.

