

Prologue

Compiler des sources

Cette partie contient :

- Compilo, compilatio, compiler
- Mosaïque
- Le plagiat est nécessaire
- Isidore de Séville pris la main dans le sac
- Virgile à travers les âges
- La compilation comme genre littéraire
- Recueil de faits survenus entre les 17 et 28 décembre 2019
- Un problème politique
- Délaissement de la compilation

Compilo, compilatio, compiler

D'après mon [Gaffiot](#), le mot latin *compilatio* vient du verbe *compilo*, qui signifie d'abord « piller, dépouiller ». « Compiler », c'est piller ses sources. C'est plagier.

Les vingt livres des *Étymologies*, composés au premier tiers du VII^e siècle dans l'Espagne wisigothique par Isidore, évêque de Séville, avaient une vocation universelle : rassembler toutes les connaissances du temps de leur auteur, ou plus précisément toutes celles qu'avaient consignées les écrivains de l'Antiquité et qui, de manuscrit en manuscrit, avaient survécu jusqu'à son temps.

Un des ouvrages les plus copiés et les plus cités au Moyen Âge et jusqu'à la Renaissance, cette encyclopédie peut être considérée comme le « chef-d'œuvre d'Isidore et de toute son époque ».

(J'ai trouvé cette formule dans un article d'Isabel Velásquez, « Les *Étymologies* d'Isidore de Séville : une vision du monde au VII^e siècle au moyen de mots », paru dans le numéro 142 de la revue *Connaissance des pères de l'Église* [Bruyères-le-Châtel, Nouvelle Cité, 2016].)

Au livre X, chapitre 44, de ses *Étymologies*, Isidore écrit :

Le *compiler* est celui qui mêle les mots d'autrui aux siens, comme les marchands de couleurs broient dans leur mortier [*pila*] différents ingrédients afin de les mélanger. Le grand poète Virgile fut un jour accusé de ce crime ; ayant glissé dans son œuvre des vers qu'il avait traduits d'Homère, et se voyant

désigné par ses rivaux comme un plagiaire [*compiler*] des Anciens, il répondit qu'il fallait être un héros pour arracher la massue des mains mêmes d'Hercule.

« Compiler », c'est donc aussi fabriquer des couleurs, c'est écrire des poèmes que les Latins admiraient – Virgile est abondamment cité dans les *Étymologies* –, et c'est se comporter en héros.

Si la présence d'un évêque espagnol du VII^e siècle, devenu docteur de l'Église au XVIII^e siècle, peut paraître incongrue dans cet écrit, je prie mon lecteur de lire ce qui suit, que je recopie depuis la [page](#) consacrée à Isidore de Séville sur Wikipedia :

À cause des *Étymologies*, dont la structure rappelle celle des bases de données et préfigure les inventions futures du classement alphabétique, puis de la notion d'index, et dont l'exhaustivité évoque le potentiel d'Internet, Isidore de Séville est proclamé [par le Vatican], en 2002, saint patron des informaticiens, des utilisateurs de l'informatique, de l'Internet et des internautes.

Mosaïque

Sur Wikipedia toujours, je trouve cette [phrase](#) de Wallace Lindsay, premier éditeur moderne des *Étymologies*, au début du XX^e siècle :

Le livre d'Isidore est une véritable mosaïque de pièces empruntées aux auteurs qui l'ont précédé, chrétiens ou païens, reproduisant souvent leurs mots mêmes sans les modifier.

Isidore's book is really a mosaic of pieces borrowed from previous writers, sacred and profane, often their ipsa verba without alteration.

Jacques Elfassi, dans un article intitulé « Isidore de Séville et les études isidoriennes aujourd'hui : une introduction » (publié dans le numéro déjà cité de la revue *Connaissance des pères de l'Église*), développe cette idée :

L'encyclopédiste, empruntant à de multiples sources, profanes comme chrétiennes, les abrège, les complète, les réécrit et surtout les réordonne selon un plan et une visée qui lui sont propres. L'art d'Isidore est comparable à celui du mosaïste : la mosaïque, si on la regarde dans le détail, est formée de multiples fragments dont aucun n'est « original » ; mais si on la regarde de plus loin, l'ensemble de ces fragments, réunis, collés les uns aux autres, forme une nouvelle œuvre, tout à fait originale.

Le plagiat est nécessaire

Je pense à un autre Isidore, nommé Ducasse, aussi connu sous le nom de Comte de Lautréamont. Au printemps 1870, quelques mois avant de mourir âgé de 24 ans, Ducasse a publié deux étranges plaquettes intitulées *Poésies* ; malgré leur titre, elles se présentent comme des suites d'aphorismes et constituent une expérimentation littéraire assez potache.

Le second volume de ces *Poésies*, qui met bout à bout et détourne des sentences des moralistes français des XVII^e et XVIII^e siècles, tels que Pascal, La Rochefoucauld ou Vauvenargues, contient ces phrases restées célèbres :

Le plagiat est nécessaire. Le progrès l'implique. Il serre de près la phrase d'un auteur, se sert de ses expressions, efface une idée fautive, la remplace par l'idée juste.

En s'appropriant les mots d'autrui, comme Virgile ceux d'Homère, comme Lautréamont ceux d'auteurs rabâchés en classe de rhétorique et de philosophie, le plagiaire perfectionne ses sources.

Je ne suis pas certain que la littérature soit soumise à des phénomènes de progrès – je préfère penser qu'elle ne tend vers rien, qu'elle change constamment, tâchant de s'adapter à son milieu et de suivre les évolutions du monde.

Il me semble en revanche que ce que Ducasse appelle « plagiat », qui consiste à utiliser ce qui a déjà été fait par un autre, tout en le corrigeant, il me semble donc que cette méthode correspond à la façon dont sont aujourd'hui développés les logiciels libres.

Un logiciel, nous le verrons, c'est d'abord une suite d'instructions, qui se présente sous la forme d'un texte écrit par un être humain et appelé « code source ». Pour qu'une machine soit capable de comprendre ces instructions, le code source doit être transformé en une suite de 0 et de 1, dite « code binaire »; l'opération permettant de passer de l'un à l'autre s'appelle la « compilation ». (Ces mots sont utilisés de même en anglais : on parle de *source code* et de *compilation*.)

Les logiciels dits libres ont cette particularité que leur code source est mis à disposition de tous, de sorte que chacun puisse l'examiner, l'améliorer – *effacer une idée fautive, la remplacer par l'idée juste* –, ou le réutiliser à sa guise.

Cette méthode a produit des résultats importants : le navigateur Web Firefox est un logiciel libre, de même que le lecteur de médias VLC, que la suite bureautique LibreOffice ou que le noyau de système d'exploitation Linux. Si la part des systèmes d'exploitation libres reste faible sur les ordinateurs individuels, Linux équipe en revanche la majorité des serveurs Web et nombre d'objets connectés, et sert de base au système d'exploitation Android, présent sur à peu près [85%](#) des smartphones. Nous reparlerons de tout cela. (Merci Éric pour m'avoir permis de préciser ces points.)

Des entreprises telles que Google, Facebook ou IBM sont désormais très impliquées dans le monde du libre ; mais le cas de Microsoft est peut-être le plus surprenant. En 2001, Steve Ballmer, son directeur général, [déclarait](#) :

En ce qui concerne la propriété intellectuelle, Linux est un cancer qui s'attache à tout ce qu'il touche.

Linux is a cancer that attaches itself in an intellectual property sense to everything it touches.

Aujourd'hui, on voit passer des nouvelles telles que celles-ci :

- en juin 2018, Microsoft rachète la plateforme [GitHub](#), qui permet aux développeurs de partager leur sources et de les éditer collectivement. GitHub est elle-même fondée sur l'outil [Git](#), développé en 2005 par Linus Torvalds, créateur de Linux ;
- en mai 2019, Microsoft [annonce](#) que Windows 10 intégrera bientôt un noyau Linux ;
- en novembre 2019, on [apprend](#) que l'une des deux personnes chargées de la maintenance des versions du noyau Linux bénéficiant d'un support à long terme est un employé de Microsoft.

On peut se réjouir de ces avancées du logiciel libre, ou au contraire méditer sur l'étonnante capacité du capitalisme à convertir en profit tout ce qui se présente comme un outil d'émancipation.

Isidore de Séville pris la main dans le sac

Je reviens à Isidore de Séville. En farfouillant sur Internet, je découvre que son anecdote concernant Virgile est empruntée à Jérôme de Stridon, et plus précisément aux quatrième et cinquième phrases de sa préface aux *Questions hébraïques* – ce *Hieronimus Stridonensis*, devenu saint Jérôme, est celui qui a réalisé, depuis le grec et l'hébreu vers le latin, la traduction de la Bible connue sous le nom de Vulgate.

Je préfère ne pas commencer à chercher en quoi consistent les *Questions hébraïques* de Jérôme, mais on peut en consulter le texte sur le site [Documenta Catholica Omnia](#).

Voici la version de Jérôme :

Hoc idem passus est ab aemulis et Mantuanus vates, ut cum quosdam versus Homeri transtulisset ad verbum, compilator veterum diceretur. Quibus ille respondit magnarum esse virium clavam Herculi extorquere de manu.

Et voici maintenant celle d'Isidore :

Hoc scelere quondam accusabatur Mantuanus ille vates, cum quosdam versus Homeri transferens suis permiscuisset et compilator veterum ab aemulis diceretur: ille respondit magnarum esse virium clavam Herculi extorquere de manu.

Ce n'est pas le sens général qu'Isidore reprend à Jérôme, mais jusqu'à ses mots. Il s'autorise toutefois deux modifications :

- Il fait entrer dans ce récit l'idée de *scelus*, de crime, qui ne figure pas chez Jérôme.
- Il écrit que Virgile a mêlé à son œuvre des vers traduits d'Homère (*quosdam versus Homeri transferens suis permiscuisset*), quand Jérôme précise qu'il les a recopiés mot pour mot (*transtulisset ad verbum*). C'est un même verbe latin que je rends ici par «recopier» et là par «traduire»: il s'agit de *transfero* (qui prend la forme *transtuli* au passé), mot qui (d'après mon [Gaffiot](#)) signifie tout à la fois «transcrire», «traduire», «transporter», «transplanter»; le plagiat et la traduction sont pour le latin deux modalités d'une même action.

Isidore se montre donc fort sévère avec la compilation, qu'il assimile à un crime; compilateur lui-même, il atténue la faute de Virgile et se plaît à citer la belle repartie du poète. Quelle ambivalence !

Voudrait-il, en étant le premier à porter cette accusation, éviter qu'elle lui soit adressée? Ou bien se pourrait-il qu'il n'eût pas conscience d'être lui-même un compilateur? Les textes que nous connaissons par cœur, assidûment fréquentés, ne finissent-ils pas par vivre en nous, par constituer une partie de nous-mêmes, par penser à notre place? Ce n'est pas Isidore qui recopie Jérôme, mais Jérôme qui parle à travers lui.

(Une question incongrue, posée en termes modernes: l'apprentissage par cœur peut-il être considéré comme une forme de copie privée, et à ce titre faire exception au droit d'auteur?)

Virgile à travers les âges

Cette anecdote à propos de Virgile est devenue un lieu commun de la littérature latine. Jérôme la tenait peut-être du grammairien [Donat](#), qui fut son professeur. Elle se lit en effet dans une vie de Virgile que rédigea Donat, restée notre principale source biographique concernant le poète.

Or, il semblerait que Donat ait lui-même, dans cet écrit, recopié une *Vita* rédigée par Suétone (qui vécut un siècle et demi après Virgile) et aujourd'hui perdue – c'est-à-dire qu'elle n'est connue que par les manuscrits de Donat, dont les spécialistes, s'appuyant eux-mêmes sur des sources antiques, estiment qu'il n'a fait que recopier l'historien romain.

Le texte de cette vie de Virgile figurait, en 1914, dans l'édition Loeb des œuvres de Suétone, dont le latin peut être lu [ici](#) et la traduction anglaise [là](#). (Merci à [Bill Thayer](#) pour les avoir mises en ligne.)

Sed hoc ipsum crimen sic defendere adsuetum ait: cur non illi quoque eadem furta temptarent? Verum intellecturos facilius esse Herculi clavam quam Homero versum subripere.

Vergil used to meet this latter accusation with these words: "Why don't my critics also attempt the same thefts? If they do, they will realize that it is easier to filch his club from Hercules than a line from Homer."

Virgile répondait ainsi à ces accusations: « Pourquoi ne s'essaient-ils pas au même larcin? Ils réaliseront alors qu'il est plus facile de s'emparer de la massue d'Hercule que des vers d'Homère. »

Je trouve la repartie de Virgile, telle qu'elle se lit chez Isidore et chez Jérôme, beaucoup plus spirituelle que chez Suétone et Donat – preuve que le plagiat réussit au personnage.

Ce prologue, que je souhaitais court, paraît déjà bien long, et il n'est pas encore achevé: je propose donc à mon lecteur que nous fassions une pause. À force de se tenir immobile devant un écran, on finit par s'abîmer les yeux et le dos.

La compilation comme genre littéraire

Comment Isidore, qui consacre un court paragraphe au *compilator*, n'a-t-il pas conscience que son œuvre-maîtresse, composée à partir de sources essentiellement livresques, est elle-même une compilation? C'est commettre un anachronisme que de poser cette question.

Dans un article intitulé « [L'Historien et la compilation au XIII^e siècle](#) », paru en 1985 dans *Le Journal des savants*, Bernard Guenée explique que c'est au Moyen Âge – au XII^e siècle plus précisément –, que le nom de *compilatio* s'est vu appliquer à ce genre littéraire, pourtant pratiqué depuis l'Antiquité, qui consistait à réunir dans un ouvrage des extraits provenant de sources diverses.

Isidore et ses contemporains ne pouvaient pas considérer les *Étymologies* comme une « compilation ». Cinq siècles plus tard pourtant, les érudits revendiquent tout à coup le titre, jusque-là infâmant, de *compilatores*.

Je cite Bernard Guenée – je me suis permis de découper son texte en paragraphes, espérant que cela faciliterait la lecture sur écran :

En s'abritant derrière ses sources, en précisant qu'il n'était pas auteur, qu'il était compilateur, nul doute que l'historien de l'*Historia Regum Francorum*, en 1214, entendait étaler sa modestie. [...] Mais plus on avance dans le XIII^e siècle, et moins est évidente la modestie qui pousse l'historien à refuser le titre d'auteur. Plus est évidente, au contraire, chez l'historien la fierté d'être compilateur.

L'étude attentive des prologues historiques du XIII^e siècle montre que, de plus en plus, l'historien justifie la compilation et chante ses vertus ; il annonce sans ambiguïté qu'il offre au lecteur une compilation ; il se nomme et s'annonce avec solennité comme un compilateur.

Au paragraphe suivant (même remarque) :

La première vertu de la compilation, qui lui est d'ailleurs reconnue depuis longtemps, c'est la brièveté. Elle offre au lecteur, sous une forme condensée, de nombreux renseignements autrement dispersés.

« Comme les faits des temps sont infinis, ceux qui les recueillent fort nombreux, et que toutes les histoires ne peuvent être ni possédées ni lues par tous, nous n'avons pas jugé inutile de retenir quelques-uns de ces faits infinis et de les réduire en un seul résumé » (*non inutile duximus ex infinitis pauca colligere et in unum coartare compendium*).

Ainsi Robert d'Auxerre commençait-il en 1181 le prologue de sa chronique. Ce prologue lui-même fut repris mot à mot, par plusieurs historiens, au XIII^e siècle, dans leur propre prologue. Et dans le même temps, d'autres historiens se déclaraient satisfaits d'avoir tiré, selon une expression consacrée, « *pauca de multis* », offrant au lecteur un commode manuel, *opus manuale, volumen manuale*.

Recueil de faits survenus entre les 17 et 28 décembre 2019

Au seuil de cet écrit que j'entends bien, si Dieu m'en donne le loisir, consacrer à l'informatique, je me retrouve à mon tour confronté à ce que Robert d'Auxerre nommait des « faits infinis ».

En voici quelques-uns que j'ai glanés depuis une dizaine de jours. (J'écris ces lignes le 28 décembre 2019.) Bien sûr, cet inventaire n'a rien d'exhaustif.

Combattre la haine en ligne

Mercredi 17 décembre. Le Sénat adopte la [proposition de loi](#) visant à lutter contre les contenus haineux sur internet (voir aussi sur [Wikipedia](#)).

Pour grande et belle et noble que soit la lutte contre la haine – on se souvient que l'entreprise Google a eu, jusqu'en 2018, cette devise: *Don't be evil* (« Ne soyez pas méchants ») –, cette loi prévoit des obligations si lourdes pour les hébergeurs de contenus, et des amendes si importantes en cas de manquement, que seules les plus grosses plateformes – telles que Youtube, Facebook, Twitter – pourront s'y conformer.

Elle aura donc deux effets :

- d'une part, favoriser ces géants au détriment de concurrents de moindre taille ;
- d'autre part, inciter les hébergeurs à pratiquer une censure des contenus de manière pro-active.

(En d'autres termes, l'État abandonne la censure à des acteurs privés.)

Concernant ce deuxième point, la Commission européenne avait adressé une [lettre](#), le 22 novembre 2019, au ministre de l'Europe et des Affaires étrangères Jean-Yves Le Drian, dans laquelle on lit :

Ceci pourrait notamment créer une charge disproportionnée sur les plateformes en ligne et, dans certaines circonstances, un risque de suppression excessive de contenus, ce qui porterait ainsi atteinte à la liberté d'expression. Ce risque est particulièrement élevé pour les plateformes en ligne dont les ressources sont limitées.

Je signale enfin que la décision de punir un hébergeur ne reviendrait pas à la justice, mais à une administration, le CSA – ce que l'association la Quadrature du Net commente [ainsi](#) :

C'est un moyen pour le gouvernement de reprendre le contrôle sur ce moyen d'expression qu'il ne maîtrise pas et cela passe par la création de lois pour et avec les grandes plateformes, en ne voyant Internet que par le prisme faussé des géants du Net, avec la menace qu'elles s'appliquent un jour à tous. Car il est toujours plus facile de ne traiter qu'avec un nombre restreint de gros acteurs (d'ailleurs plus prompts que les autres à collaborer avec lui).

Le texte de la loi sera examiné le 8 janvier par une [commission mixte paritaire](#).

J'ouvre une parenthèse.

On trouve sur le site de la Haute Autorité pour la Transparence de la Vie Publique des informations concernant l'activité de lobbying de [Google](#), propriétaire de Youtube, en France.

Les chiffres pour 2019 n'ont pas été publiés (la date limite de déclaration étant le 30 mars) mais, en 2018, l'entreprise a dépensé entre 1 et 1,25 million d'euros pour la défense de ses intérêts.

On peut voir [ici](#) que son activité a concerné (entre autres) l'avant-projet de la loi contre la haine, l'entreprise s'étant livrée à une « présentation des mesures de détection des contenus, de traitement des signalements et des actions prises contre les auteurs ».

Le projet de loi ayant été dévoilé au printemps 2019, puis examiné par l'Assemblée nationale en juillet de la même année, il est fort probable que d'autres actions ont été menées par Google.

J'en profite pour signaler que Google effectue aussi un intense travail de lobbying en Europe. L'entreprise est celle qui a obtenu le plus de rendez-vous auprès des membres de la Commission européenne depuis 2014 – voir sur le site [Integrity Watch](#), maintenu par l'association Transparency International.

Fin de la parenthèse.

Combattre la fraude fiscale

Jeudi 19 décembre. L'assemblée nationale adopte la [loi de finance pour 2020](#), dont l'article 154 prévoit que, « à titre expérimental et pour une durée de trois ans », l'administration fiscale et celle des douanes pourront « collecter et exploiter au moyen de traitements informatisés et automatisés n'utilisant aucun système de reconnaissance faciale les contenus, librement accessibles sur les sites internet des opérateurs de plateforme en ligne [...] manifestement rendus publics par leurs utilisateurs ».

Si je comprends bien, il s'agit, afin de lutter contre la fraude fiscale, de récupérer ce que les citoyens français postent sur les réseaux sociaux ainsi que sur les sites de vente en ligne tels que le Bon Coin.

Quand le gouvernement, au mois de septembre dernier, a présenté son [projet](#) de loi – qui était défini par cette belle formule : « Le projet de loi de finances pour 2020 consacre l'acte 2 du quinquennat, dont il traduit les ambitions politiques : baisser massivement les impôts et préparer l'avenir » –, les mots « collecter en masse » apparaissaient dans le texte ; ils en ont aujourd'hui disparu.

Collecter des données de géolocalisation

Jeudi 19 décembre (également). Le *New York Times* [annonce](#) qu'il s'est procuré une base de données provenant d'une entreprise spécialisée dans la géolocalisation.

La liste contient plus de 50 milliards de signaux envoyés par les ordiphones de plus de 12 millions d'individus en déplacement dans quelques grandes villes des États-Unis – soit une

moyenne de 4166 signaux par appareil. La collecte s'étale sur une période de plusieurs mois (sans autre précision) en 2016 et 2017.

On trouve sur [cette image](#) un aperçu du fichier concerné, un simple [fichier en texte brut](#) (ici avec l'extension [.csv](#)), qui peut être ouvert et édité avec n'importe quel éditeur de texte (merci encore Éric pour ces précisions!).

Les auteurs de l'article expliquent qu'en fouillant dans ces données, ils ont pu savoir qui a rendu visite à Johnny Depp, Arnold Schwarzenegger ou Tiger Woods; qu'ils ont deviné qu'un employé de Microsoft allait passer un entretien d'embauche chez Amazon; qu'ils ont retrouvé la trace d'un haut fonctionnaire du ministère de la Défense et de son épouse au sein de la Marche des Femmes, organisée au lendemain de l'investiture de Donald Trump; qu'ils ont suivi celle de journalistes – il est donc possible, par recoupement, de savoir qui un journaliste rencontre, et donc, possiblement, de connaître ses sources.

Identifier des visages

Jeudi 19 décembre (toujours). Une énième [étude](#) sur la reconnaissance faciale, cette fois-ci publiée par le National Institute of Standards for Technology, insiste sur les biais de cette technologie.

L'enquête a testé 189 algorithmes dans deux tâches :

- La première consiste à trouver, pour deux photographies données, si elle représentent la même personne;
- La seconde à trouver, pour une photographie donnée, si d'autres photographies de la même personne se trouvent dans un jeu de photographies.

Les 18 millions d'images utilisées dans cette étude proviennent des bases du FBI, du département de la Sécurité intérieure et du ministère des Affaires étrangères; elles représentent près de 8,5 millions d'individus.

Il apparaît que les faux positifs (cas où l'algorithme identifie les photos de deux personnes différentes comme étant celles d'une même personne) sont de deux à cinq fois plus élevés – le chiffre varie selon les algorithmes – pour les femmes que pour les hommes; ils sont de 10 à 100 fois plus élevés pour les Asiatiques et les Afro-américains que pour les Caucasiens (ah qu'en termes galants ces choses-là sont dites). En résumé: les femmes et les noirs courent un risque très élevé d'erreur judiciaire.

Alors que nous nous apprêtons à entrer dans l'année 2020, je tiens toutefois à adresser un message d'espoir à ces algorithmes si durement jugés par cette étude. Car moi non plus, je ne suis pas physionomiste; je voulais qu'ils sachent qu'on peut très bien vivre avec cela.

Je hasarde enfin cette idée: la vraie question, la seule question importante, ce n'est pas celle du degré d'exactitude des techniques de reconnaissance faciale; c'est de savoir s'il est possible d'accepter que ces techniques soient utilisées dans l'espace public, dans notre quotidien, dans nos vies. Peut-être réussirait-on, en arguant de la médiocrité de leurs performances, retarder leur déploiement; mais je suppose que les progrès dans ce domaine seront rapides.

Suivre chacun dans ses déplacements

Samedi 21 décembre. Le *Journal officiel* publie un « [arrêté ministériel](#) du 16 décembre portant création d'un service à compétence nationale dénommé "Service national des données de voyage" »:

Le Service national de données de voyage est chargé de la mise en place et de l'amélioration des dispositifs de collecte et d'exploitation des données de voyage, à savoir les données de réservation, d'enregistrement et d'embarquement des passagers, et le cas échéant des équipages, des transports aériens, maritimes et terrestres [...].

Interrogée par le journaliste Xavier Berne, de [Next INpact](#), la Direction générale de la police nationale résume ainsi le rôle de ce nouveau service :

L'idée, c'est de tracer les trajets des passagers. Au départ, avec le PNR, ce n'était que l'avion. Désormais, on l'étend au maritime et au ferroviaire.

Expérimenter

Mardi 24 décembre. Ayant probablement négligé de lire l'étude publiée le 19 par le National Institute of Standards for Technology, Cédric O, secrétaire d'État chargé du numérique, déclare au [Parisien](#) :

Le gouvernement souhaite donc ouvrir une phase d'expérimentation, de six mois à un an, sous la supervision de la société civile et des chercheurs. Cela doit nous permettre d'avoir des retours sur les différents cas d'usage et sur la performance réelle de cette technologie, ainsi que sur les questions qui se posent.

Aux [Échos](#), son cabinet déclare que « seuls les individus qui auront donné leur consentement seront identifiés lors de cette expérimentation ». Nous voilà rassurés!

Plutôt que d'expérimenter, je suppose qu'il s'agit de faire savoir à la population qu'elle devra s'habituer à être reconnue où qu'elle aille.

(La déclaration de Cédric O a été faite au cours d'un entretien surtout consacré à [Alicem](#), « solution d'identité numérique » développée par l'Agence nationale des titres sécurisés, associant passeport biométrique, reconnaissance faciale et application tournant sur Android).

S'ouvrir à d'autres cultures

Jedi 26 décembre. À peine remis des trois douzaines d'huîtres que j'ai avalées sur vingt-quatre heures glissantes, je lis dans un [communiqué de presse](#) que la ville de Brienon-sur-Armançon, 3300 habitants, située en Bourgogne, a « adopté la solution Safe City » développée par l'entreprise chinoise Dahua Technology.

Creusant un peu, j'apprends [ici](#) que Dahua Technology joue un rôle clé dans la surveillance de masse des Ouighours, victimes d'une persécution ethnico-religieuse rendue possible par l'intelligence artificielle (*mass surveillance of Uyghurs under AI-enabled ethno-*

religious persecution) et [là](#) que le gouvernement américain, en protestation contre le traitement réservé par Pékin aux Ouïghours, a récemment placé l'entreprise sur une liste noire.

Mais revenons à Briennon-sur-Armançon, dont le communiqué de presse nous assure qu'elle « pâtit [...] d'une forte activité criminelle » et proclame fièrement :

Durant la première semaine d'essai de ce système intelligent, la police a déjà résolu une affaire de vol dans un cimetière grâce à des preuves enregistrées par la caméra PTZ de Dahua.

(Il n'est donc plus permis d'aller pleurer ses morts sans être observé ; ceci dit, au cas où les morts finiraient par se réveiller, la présence de ces caméras de *vidéoprotection* – puisque l'[article 17](#) de la loi du 14 mars 2011 d'orientation et de programmation pour la performance de la sécurité intérieure stipule que « sous réserve des dispositions de la présente loi, dans tous les textes législatifs et réglementaires, le mot : “vidéosurveillance” est remplacé par le mot : “vidéoprotection” » – ne pourra que renforcer notre sécurité.)

Un [article](#) paru le 9 décembre dans *L'Yonne républicaine* nous informe que la commune est désormais équipée de 29 caméras de surveillance (soit une caméra pour 113 habitants), qui permettent d'assister l'unique policier municipal dans sa mission :

Briennon-sur-Armançon a déboursé 52 000 € pour cette installation. Un accord entre la Ville et Dahua Technology a permis de baisser le prix en échange d'une promotion des produits de cette dernière auprès des élus locaux.

(Une douzaine d'élus des communautés voisines ont assisté, le 7 décembre, à la réunion de présentation du dispositif.)

On apprend aussi que les caméras de surveillance peuvent lire de 95 à 98% des plaques minéralogiques des véhicules entrants et sortants.

Continuer le combat contre la fraude fiscale

Vendredi 27 décembre. Le Conseil constitutionnel, qui avait été saisi au sujet de la loi de finance 2020, la valide. Je lis cela sur [Next INpact](#) :

Dans sa décision, le Conseil constitutionnel relève que Big Brother Bercy permettra « d'une part, de collecter de façon indifférenciée d'importants volumes de données, relatives à un grand nombre de personnes, publiées sur de tels sites et, d'autre part, d'exploiter ces données, en les agrégeant et en opérant des recoupements et des corrélations entre elles ».

Même si les données sont rendues publiques par les personnes concernées, il y a atteinte à la vie privée. Et « dans la mesure où elles sont susceptibles de dissuader d'utiliser de tels services ou de conduire à en limiter l'utilisation, [les dispositions contestées] portent également atteinte à l'exercice de la liberté d'expression et de communication ».

Cependant, ceci posé, le conseil considère que l'atteinte n'est pas déséquilibrée, ou insupportable face à l'objectif de lutte contre les infractions fiscales et douanières.

En somme : il y a collecte indifférenciée d'importants volumes de données et atteinte à la liberté d'expression et de communication, mais il faut savoir ce qu'on veut.

Retrouver les enfants disparus

Samedi 28 décembre. Dans une [dépêche](#) publiée sur le site Bloomberg.com, je lis qu'un logiciel de reconnaissance faciale, déployé en 2018 en Inde dans le but d'aider la police à retrouver des enfants disparus, est désormais utilisé pour identifier les participants à des manifestations hostiles à une réforme de la citoyenneté [âprement combattue](#).

Un problème politique

Chacune de ces nouvelles semble bien ordinaire, insignifiante, pour ne pas dire assommante – le gouvernement, les algorithmes, les Chinois, les Indiens, on sait très bien à quoi s'en tenir sur ces sujets ; inutile d'insister.

Des faits passent ainsi chaque jour sans qu'on les remarque ; juxtaposés, ils apparaissent comme autant d'atteintes, plus ou moins discrètes, plus ou moins subtiles, à la vie privée, au secret, enfin à la liberté. Il faut, pour les enregistrer toutes, faire preuve d'une vigilance continuée, qui épuise. La tâche est impossible.

Je ne prétends pas que les ordinateurs – puisque, malgré les apparences, c'est bel et bien d'eux que je souhaite entretenir mon lecteur – soient la seule source de ces atteintes. Il se trouve pourtant que nous ne savons pas comment fonctionnent ces machines ; que nous connaissons mal les possibilités qu'elles offrent à ceux qui les contrôlent ; que nous devinons seulement ce que permettent leurs capacités de collecte, de traitement et de stockage de l'information.

L'ignorance dans laquelle nous nous maintenons au sujet de ces appareils, qui régissent nos vies, qui équipent le moindre des objets que nous manipulons au quotidien, a de quoi surprendre.

Au cours de ces recherches, je me suis souvent fait ces deux réflexions :

- La première, c'est qu'en informatique la réalité dépasse, ou du moins précède, toujours notre imagination. De tout ce que j'ai imaginé de déplaisant, d'inquiétant ou de surprenant qu'un ordinateur permette de faire, il n'est rien dont je n'ai pas trouvé qu'il le fasse déjà.
- La seconde, c'est que la plupart des problèmes qu'on attribue à l'informatique comme à une sorte de fatalité ou de volonté supérieure – « c'est la faute aux ordinateurs, aux algorithmes, etc. » – sont politiques ; et par ce mot de politique, je ne vise pas les hommes et les femmes qui nous dirigent, mais nous-mêmes. Nos aspirations, nos comportements, nos choix ne sont peut-être pas compatibles avec la vie privée et la liberté.

Je m'aperçois qu'il reste quelques petites choses à dire à propos de la compilation. Avançons donc.

Délaissement de la compilation

Avec la Renaissance, puis les Lumières, la compilation perd de son prestige.

Arrachées aux érudits, à leurs mains tâchées d'encre, à leur myopie, les études historiques passent sous le contrôle de penseurs visionnaires et puissants, de personnalités fortes capables d'imprimer leur marque aux événements du monde.

Plutôt que d'établir patiemment, à partir de sources péniblement collationnées et souvent contradictoires, un récit qu'on sait imparfait, lacunaire, parcellaire, plutôt que de décrire minutieusement, humblement un monde obscur, dont la connaissance se refuse à nous, il s'agit désormais de soumettre en esprit cet objet qu'on nomme *réalité*, de le réduire en mots, en idées, en causes, en conséquences, en périodes, en époques, en théories – puis bientôt en fichiers, en bases de données, en 0 et en 1 –, de le dominer, de le vaincre.

Délaissée, la compilation devient le fait d'auteurs sans originalité, sans pensée, sans éclat. En 1754, dans le volume III de l'*Encyclopédie*, l'Abbé Mallet la [définit](#) comme un « recueil formé de morceaux pris çà et là dans le même ou dans divers auteurs »; il évoque des « copies sèches ou informes de lambeaux mal cousus », un « amas de matériaux bruts ».

Et s'étant faite silve, trésor, miroir, spicilège, miscellanées, la compilation de là dégénère en ana, en varia, en *Paroles du Bouddha*, en *Marcel Proust en verve*, en extraits du *Readers's Digest*, en *Apprenez l'histoire de France aux toilettes*.

J'ai placé là ces deux dernières références à seule fin de rappeler que Robert d'Auxerre, dans la première phrase de son prologue [*cum infinita sint temporum gesta gestorumque digestores quam plurimi*] nomme *digestores* les auteurs de compilations. Je retourne à mon [Gaffiot](#), qui m'apprend qu'un *digestor* est « un ordonnateur, un organisateur »; que le verbe *digero* signifie « diviser, séparer, distribuer, répartir, mettre en ordre, classer »; que la *digestio*, c'est la « répartition de la nourriture dans le corps »; qu'une *digestio Italiae in litteras*, c'est une « description méthodique de l'Italie »; et enfin que les *Digesta*, désignent « toute espèce d'œuvre distribuée en chapitres, en particulier les livres des Pandectes », qui sont pour leur part des recueils de décisions prises par des juriconsultes romains. Un écrivain, c'est une [bête qui digère](#).

Mais le temps est enfin venu de parler d'informatique. J'ai obéi, en composant ce ce qui suit, autant à mon incompetence qu'à ma fantaisie. Si l'ensemble paraît disparate, pour ne pas dire qu'il frappe par son incohérence, du moins présente-t-il cet avantage que ses différentes parties peuvent se lire indépendamment les unes des autres. Chacun piochera selon ses envies.



Ce texte est publié sous licence Creative Commons.