

PARADOXALE

LA REVUE



NUMÉRO
10

ISSN 2981-2771

MENSUELLE

OCTOBRE 2026

Le piège du flacon

Chronique d'un naufrage numérique

octobre 2026
collection « PARADOXALE | LE LAB DES CONTRE-SENS »

Diffusé sous licence Creative Commons
Attribution – Pas d'Utilisation Commerciale – Pas de Modification
4.0 International (CC BY-NC-ND 4.0)
creativecommons.org/licenses/by-nc-nd/4.0/deed.fr

Outre les droits accordés par cette licence, l'auteur autorise
expressément la reproduction et la diffusion interne de ce numéro
au sein des entreprises, administrations et établissements
d'enseignement, ainsi que son usage à des fins de formation,
y compris lorsque ces organisations poursuivent un but lucratif.

Le droit moral de l'auteur demeure inaliénable.
© 2026 Sébastien VANNEROT.

ISSN 2981-2771 (en ligne)

Table des matières

Éditorial	7
Le piège du flacon	9
Introduction	9
1 La doctrine du contenant numérique	12
1.1 Répondre du contenu et du contenant	12
1.2 Stocker n'est pas conserver	14
1.3 Le choix du réglage par défaut	18
1.4 Sous la ligne de flottaison	20
2 Les étages de la chaîne numérique	22
2.1 La fausse nudité de la donnée	22
2.2 Le format de la traversée	23
2.3 Le gabarit du logiciel	27
2.4 La carte du sens	28
2.5 La loi de l'hébergeur	30
2.6 Les détroits de la diffusion	31
2.7 La persistance des défauts	33
2.8 La licence égarée	33
3 Une loi de surface	35
3.1 Le contrôle du pavillon	35
3.2 Le format, entre vœu et lettre morte	37
3.3 Le choix allemand	39
3.4 L'interopérabilité en question	40
3.5 Le droit de reprendre sa cargaison	41
3.6 Un code public sans port d'attache	43
3.7 La licence européenne en suspens	43
3.8 Modèles et plateformes : le silence des textes	44
3.9 Une mesure sans conséquence	44
3.10 Faites ce que je dis, pas ce que je fais	45
3.11 Une digue de papier	46
4 Le régulateur face à ses choix	48
4.1 Le droit et la machine	48
4.2 Un modèle venu de Chine	49
4.3 Un coût de sortie jamais évalué	50

4.4	Une dépendance active	51
4.5	Entre Pékin et Washington.....	51
5	La pérennité du contenant	53
5.1	La lisibilité de la preuve	53
5.2	La dépossession de fait	54
5.3	L'archivage, la douane du port d'arrivée	55
5.4	Le coffre et la clef.....	56
	Conclusion.....	58
	Sources et références	63
	L'auteur	71

Éditorial

Ce numéro est né d'une contrariété plus que d'une lecture. Je travaille au développement d'un portail d'intelligence territoriale, et ce projet partait d'une intention simple, presque naïve dans sa formulation : bâtir un outil qui ne dépende de personne. L'intention a résisté à peu de choses. À chaque étape, un choix que je croyais technique se révélait être un choix de dépendance : le format d'un fichier, l'encodage qui décide si ses accents survivront au voyage, la plateforme par laquelle il circule, le modèle d'intelligence artificielle qui permet d'y chercher. Aucune de ces dépendances n'avait vraiment été choisie par quiconque. Chacune était déjà en place, et il fallait composer avec elle.

L'article qui suit est né de ce heurt entre une intention d'indépendance et la chaîne réelle du numérique. Il en garde la marque. Plus concret que les numéros précédents, et je crois plus accessible, il part d'objets que chacun manipule sans y penser : un fichier, une extension, un SMS, une recherche sur Google. Ce n'est pas une concession à la vulgarisation. Le sujet l'exige : le contenant numérique ne se laisse pas saisir dans l'abstrait, il se constate le jour où un fichier refuse de s'ouvrir.

Le titre détourne un vers que Musset place dans la dédicace de *La Coupe et les Lèvres* : « Qu'importe le flacon, pourvu qu'on ait l'ivresse ? » Chez le poète, la formule suit une autre question, « qu'importe la maîtresse ? », et plaide pour le sentiment contre son objet. En régime numérique, elle devient un piège. Le flacon, c'est le contenant : le format, l'encodage, le logiciel, le modèle, la plateforme. L'ivresse, c'est le contenu, ce que l'on veut dire, montrer ou transmettre. Croire que seul le contenu compte est une erreur d'autant plus confortable qu'elle ne se voit pas le jour de la publication, où tout semble fonctionner. C'est pourtant la forme numérique qui décide si le fond parviendra, lisible, jusqu'à son lecteur.

Le sous-titre ne prophétise rien. Il rapporte ce qui est déjà arrivé. Des œuvres entières conçues pour Flash ne s'ouvrent plus qu'à travers des programmes d'imitation entretenus par des bénévoles. Une étude publiée en 2024 a établi que 38 % des pages web qui existaient en 2013 étaient devenues inaccessibles dix ans plus tard. Sur la plateforme nationale des données publiques, des fichiers sont défigurés dès leur première ligne. Le sous-titre relève donc de la chronique, non de l'augure.

On ne trouvera dans ces pages ni coupable ni procès. L'agent a exporté son fichier comme son logiciel le lui proposait, le technicien a retenu la solution que ses critères lui désignaient, la norme a réglé ce qu'elle voyait. Chaque décision se défend ; c'est leur

accumulation qui produit un résultat que personne n'a souhaité. L'exercice ne m'a pas rendu indépendant. Il m'a appris à savoir de quoi je dépends et à me demander, pour chaque dépendance, ce que coûterait d'en sortir. C'est cette question que ce numéro voudrait remettre à ceux qui décident : non pas quel flacon choisir, mais qui l'a choisi pour eux.

Sébastien VANNEROT

Directeur de la publication

Le piège du flacon :

Chronique d'un naufrage numérique

Introduction

En 2017, un arrêté pris au nom du Premier ministre a fixé la manière dont les administrations doivent publier, en données ouvertes, c'est-à-dire librement accessibles et réutilisables par tous, les informations relatives aux subventions qu'elles accordent. La première ligne de son annexe est simple : les fichiers doivent être encodés en UTF-8, abréviation de l'anglais *Universal Character Set Transformation Format, 8 bits*. En effet, un ordinateur ne connaît ni lettre ni alphabet. Il ne connaît que des nombres et doit disposer d'une table de correspondance pour savoir que tel nombre désigne un « é » et tel autre un « ç ». Cette table s'appelle un encodage.

La première table largement adoptée, l'ASCII (*American Standard Code for Information Interchange*), date de 1963. Elle ne comptait que 128 signes : assez pour écrire l'anglais, mais sans aucune lettre accentuée. Chaque pays, puis chaque fabricant, l'a alors complétée à sa manière, si bien qu'un même nombre désignait un « é » sur un ordinateur et un tout autre signe sur l'ordinateur voisin. Pour sortir de cette confusion, un consortium d'industriels a entrepris à partir de 1991 de recenser dans un répertoire unique, l'Unicode, tous les caractères de toutes les écritures du monde. La version publiée en septembre 2025 en compte 159 801, de l'alphabet latin aux idéogrammes chinois, en passant par l'arabe, le cyrillique et les émojis. L'UTF-8 est la manière la plus répandue d'inscrire ces caractères dans un fichier : c'est la table universelle, celle que tout logiciel moderne sait lire. L'imposer, c'est garantir qu'un fichier publié à Paris s'affichera correctement à Lyon, à Bruxelles ou à Montréal, mais aussi à Tokyo ou à Abou Dhabi.

Sur la plateforme nationale des données publiques, data.gouv.fr, les services du Premier ministre publient les subventions qu'ils ont accordées de 2017 à 2021. Le jeu de données, publié par une organisation que la plateforme certifie comme service public, se réclame du modèle national conçu pour appliquer cet arrêté et a été téléchargé plusieurs milliers de fois. Pourtant, aucun des cinq fichiers annuels n'est écrit dans l'encodage requis. Dès la première ligne, l'intitulé de colonne « Référence de la décision », que l'arrêté a lui-même défini, apparaît déformé pour tout logiciel qui attend la table universelle. En outre, le fichier ne porte aucune licence, ce document qui dirait au lecteur ce qu'il a le droit d'en faire. Et le programme de contrôle automatique de la plateforme, qui l'a examiné au début

de 2026, l'a classé comme « texte, sans erreur ».

Aucun des acteurs de cette affaire n'a pourtant commis de faute; chacun n'a fait qu'un geste anodin. L'agent a exporté le fichier comme son logiciel le lui proposait. La norme existait et elle était précise, même si l'on verra qu'elle comportait sa propre imperfection. Le programme de contrôle a lu le fichier avec les moyens dont il dispose, pour son propre usage, et l'a trouvé lisible. Dans le tableur réglé en français avec lequel il a été produit, le fichier s'ouvre vraisemblablement sans difficulté : il n'est illisible que pour un logiciel réglé autrement. Aucun de ces gestes, pris isolément, n'est donc critiquable; c'est leur addition qui produit un fichier illisible. Le paradoxe est d'autant plus net que la règle et le fichier qui l'enfreint émanent de la même autorité : les services du Premier ministre.

Le cas n'est ni isolé ni grave en soi. Sur un échantillon de fichiers publics mesuré en septembre 2026, ceux qui ne respectaient pas l'encodage attendu allaient des services du Premier ministre à une petite commune et couvraient une dizaine d'années de publications. Si cet exemple anodin a été retenu pour ouvrir cet article, c'est qu'en réalité il réunit, dans un seul fichier de quelques lignes, tous les ingrédients du problème.

Ces ingrédients sont au nombre de trois. Les chiffres du fichier ne sont pas en cause : c'est leur contenant, l'emballage numérique qui les transporte, qui les rend illisibles. Une règle existait pourtant pour prévenir ce défaut, et un contrôle automatique a bel et bien examiné le fichier. Mais ce contrôle ne vérifiait pas ce que la règle exigeait, et personne n'avait pour mission de rapprocher la règle, le contrôle et le fichier.

Chacun connaît d'ailleurs ce problème sans lui avoir jamais donné de nom. Le courriel où « é » devient « Ã© ». Le document Word qui perd toute sa mise en page lorsqu'on l'ouvre avec un autre logiciel. La présentation dont les polices changent sur l'ordinateur de la salle de réunion, au pire moment. Les vieux fichiers d'un logiciel disparu, retrouvés au fond d'un disque dur, et que plus personne ne sait ouvrir. La photographie prise avec un iPhone qui refuse de s'afficher sur l'ordinateur d'un collègue, parce que le téléphone l'a enregistrée, par défaut, dans un format que certains ordinateurs ne savent pas lire. Le document rédigé sur un Mac avec le traitement de texte d'Apple, que personne ne sait ouvrir sur un PC. L'utilisateur de Linux ou le propriétaire d'un téléphone Android à qui l'on envoie une pièce jointe que son logiciel affiche de travers. Personne n'échappe à cette expérience, quel que soit le camp technologique auquel il appartient. Dans chacun de ces cas, le contenu est intact; seul son contenant fait défaut.

Les réunions entre services de l'État et entreprises ont parfois donné lieu à une scène plus comique. L'entreprise a préparé sa présentation sous PowerPoint; l'administration la projette avec la suite bureautique libre installée sur ses postes. Les titres débordent, les images se chevauchent, les polices changent, les schémas soigneusement alignés partent

dans tous les sens. La salle sourit et l'orateur s'excuse ; la réunion reprend. Personne n'a commis de faute : chacun a travaillé avec l'outil qu'on lui avait donné. Mais deux flottes se sont croisées en pleine séance officielle, et aucune n'était construite pour naviguer dans les eaux de l'autre.

Ces désagréments font sourire tant qu'ils restent privés. Ils changent de nature lorsque le document n'est plus un courrier entre collègues, mais une publication qui engage une organisation : une donnée publique, un bilan, un acte, une pièce remise à un juge. La question devient alors celle de la responsabilité.

De quoi un dirigeant, public ou privé, répond-il lorsque son organisation publie quelque chose, un texte, un bilan, des données, un document technique ? De ce qu'il dit, sans aucun doute. Mais aussi, soutiendra-t-on, de la forme dans laquelle il le dit. Cette forme n'est pas ici le style. C'est l'objet numérique qui porte le texte jusqu'à son lecteur. Il comprend le fichier, avec son encodage et son format, le logiciel qui l'a produit, le lieu où il est stocké, la voie par laquelle il circule et la licence qui en fixe l'usage. Cet article l'appelle le contenant, par opposition au contenu qu'il transporte. C'est en effet ce contenant qui décide de ce qui restera lisible, retrouvable et transmissible, aujourd'hui, dans dix ans, dans cinquante ans, dans un siècle.

Pour l'expliquer, on filera tout au long de ce texte une métaphore maritime, celle d'un navire. La donnée est la cargaison : elle peut être de n'importe quel type, un texte, un tableau de chiffres, une image, un film. Le fichier qui la porte est la coque du navire.

Le fichier des subventions évoqué en début d'introduction est une péniche : simple et économique, elle est parfaitement adaptée aux voies d'eau intérieures, mais s'échoue dès que la voie n'a plus la profondeur qu'exige son tirant d'eau. Le tirant d'eau est la profondeur à laquelle la coque s'enfonce dans l'eau. Hors de la métaphore, le tirant d'eau est l'encodage du fichier et la voie d'eau le réglage du logiciel qui l'ouvre : lisible dans un tableur réglé comme celui qui l'a produit, le fichier s'échoue dans tous les autres. On verra au fil de ce numéro que la même traversée se joue, à une tout autre échelle, au cœur même de l'État. Or le droit, qui inspecte si soigneusement la nationalité des navires (le pavillon, dans notre métaphore), ne regarde presque jamais leur coque (le fichier lui-même, première pièce du contenant).

I La doctrine du contenant numérique

I.1 Répondre du contenu et du contenant

Personne ne conteste qu'un dirigeant doive répondre de l'exactitude de ce qu'il publie. Qu'une administration mette en ligne une donnée fausse, qu'une entreprise remette une pièce erronée à son commissaire aux comptes et chacun sait quelle responsabilité rechercher. Les questions sont connues : qui a validé, sur la foi de quelle source, avec quel contrôle ? La responsabilité du contenu est une évidence partagée. Le droit l'organise et la déontologie la rappelle ; l'opinion la réclame, si bien que cet article n'a rien à y ajouter.

Pourtant, l'évidence cesse là où commence le contenant. Rien n'indique que les montants ou les dates du fichier des subventions soient inexacts : ce qui est en cause, c'est la manière dont le fichier les porte. Trois éléments de ce contenant suffisent à le montrer, et chacun les a rencontrés sans le savoir. L'encodage, cette table de correspondance entre nombres et lettres (l'ASCII hier, l'UTF-8 aujourd'hui), décide si les caractères arriveront tels qu'ils ont été saisis. C'est lui qui, mal accordé entre l'expéditeur et le destinataire, transforme un « é » en « Ã© » dans un courriel. Le format, que signale l'extension placée à la fin du nom du fichier (.docx, .pdf, .csv), décide quel logiciel saura l'ouvrir et comment il le restituera. C'est lui qui fait déborder les titres d'une présentation projetée sur un autre ordinateur. La licence, enfin, est le document juridique qui accompagne le fichier et dit ce que chacun a le droit d'en faire : le copier, le modifier, le réutiliser, y compris à des fins commerciales. Son absence laisse l'utilisateur scrupuleux dans l'incertitude.

Autrement dit, un contenu peut être parfaitement exact et ne jamais parvenir à son lecteur dans un état utilisable. Ce qui arrive entre ses mains dépend entièrement du contenant. Celui-ci ne se réduit d'ailleurs pas au fichier. Entre la donnée saisie et l'écran qui l'affiche, le contenu traverse une chaîne entière, que l'on décrira plus loin étage par étage. Outre la donnée elle-même, elle compte six étages. Ce sont le format du fichier et son encodage, le logiciel qui le produit, le modèle d'intelligence artificielle qui permet désormais d'y chercher, le lieu où il est hébergé, la plateforme par laquelle il est diffusé et l'appareil sur lequel il est lu. La licence, qui dit ce que chacun peut en faire, l'accompagne d'un bout à l'autre.

Cette dépendance ne se voit pas le jour de la publication, où tout semble fonctionner. Le temps seul la révèle. Le critère qui départage deux contenants est donc leur durée de vie : le fichier pourra-t-il encore être lu et exploité dans cinq ans ou dans vingt ? Son lecteur n'aura alors ni le même logiciel, ni les mêmes réglages, ni personne à qui demander comment il avait été produit.

La question qui fonde cet article en découle. Si le contenant décide de ce qui restera lisible,

retrouvable et transmissible, le dirigeant peut-il s'en désintéresser au motif qu'il s'agit d'un choix technique, laissé aux services informatiques ? La thèse soutenue ici est qu'il ne le peut pas : il répond autant du contenant que du contenu. Elle se heurte pourtant à deux objections sérieuses. Aucun dirigeant ne choisit lui-même le format de chaque fichier ; et le droit, on le verra, ne le lui demande guère. La suite de cette partie y répondra, après avoir précisé ce que recouvrent les trois qualités en jeu.

Ces trois termes, lisible, retrievable et transmissible, ne sont pas synonymes. Lisible désigne la capacité d'ouvrir le document et d'en restituer fidèlement les signes : une lettre écrite d'une encre qui pâlit est intacte le jour de l'envoi, mais illisible quelques années plus tard. Retrouvable désigne la capacité de repérer un document, et ce qu'il contient, parmi tout le reste. Elle suppose que le document ait été indexé, décrit et classé de façon qu'un catalogue ou un moteur de recherche puisse le retrouver ; mais l'index ne suffit pas, car ce qu'il signale doit encore pouvoir être ouvert. Une bibliothèque dont le catalogue a brûlé conserve l'intégralité de son fonds, mais plus personne n'y trouve rien. Transmissible désigne la capacité de remettre le document à un tiers, personne ou machine, sans qu'il perde en route une part de lui-même. Un colis que le destinataire ne peut ouvrir que s'il possède la même clef que l'expéditeur n'est pas réellement délivré.

Ces trois qualités tiennent donc toutes à ce qui porte le contenu : le contenant.

Le droit de la preuve offre une illustration de cette distinction entre contenu et contenant. La signature électronique, que le Code civil a consacrée, garantit deux choses : l'identité de celui qui signe et le fait que rien n'a changé depuis la signature. Elle ne garantit rien d'autre. Un fichier vide peut être signé. Un fichier dont les caractères sont déjà illisibles peut l'être tout autant, de même qu'un fichier dans un format qu'aucun logiciel ne saura plus ouvrir. Dans tous ces cas, la signature remplira fidèlement son office : elle certifiera intact ce qui était déjà vide, altéré ou illisible. Elle scelle l'objet tel qu'il a été remis, mais ne garantit pas qu'il pourra être lu. L'instrument de confiance le plus élaboré que le droit ait construit pour l'écrit numérique porte ainsi sur le contenu et sur son auteur ; le contenant lui échappe entièrement.

Pour autant, cette distinction ne déplace pas la responsabilité de l'un vers l'autre. Présenter le contenu et le contenant comme une alternative, comme si le dirigeant devait répondre de l'un à l'exclusion de l'autre, serait poser un faux dilemme. Le renverser suffit à s'en convaincre : aucun dirigeant ne soutiendrait que le soin apporté à la forme dispense de l'exactitude du fond. La réciproque n'est pas moins vraie, mais la pratique l'oublie. Les deux responsabilités ne s'excluent pas, elles se cumulent. Elles portent sur des objets distincts et servent une même fin : que le document parvienne à son destinataire, lisible et utilisable, aujourd'hui comme demain.

Le droit appréhende ce cumul depuis déjà un siècle, pour un autre contenant. La convention conclue à Bruxelles en 1924 sur le transport de marchandises par mer distingue, pour le transporteur, deux obligations. La première porte sur le navire : le transporteur doit exercer une diligence raisonnable pour le mettre en état de naviguer et apprêter les cales qui recevront les marchandises. La convention situe cette obligation « avant et au début du voyage ». La formule signifie qu'elle s'apprécie au moment du départ, de la préparation du navire jusqu'à son appareillage, et non tout au long de la traversée. Le transporteur répond de l'état dans lequel il a fait partir son navire ; l'obligation de navigabilité ne s'étend pas à ce qui survient ensuite en mer, qui relève d'autres règles. La seconde obligation porte sur la cargaison : le transporteur doit procéder avec soin au chargement et à la garde des marchandises. L'une ne se confond pas avec l'autre, et le respect de l'une ne dispense pas du respect de l'autre. Le transporteur qui a pris soin de chaque colis n'est pas quitte s'il les a chargés dans la coque d'un navire qui prend l'eau.

Le droit français a repris ce principe. Le Code des transports impose au transporteur de « faire diligence » pour mettre le navire en état de navigabilité « compte tenu du voyage qu'il doit effectuer et des marchandises qu'il doit transporter ». La navigabilité n'est donc pas une qualité absolue. Un même navire peut être apte à une traversée et totalement inapte à une autre, apte à une cargaison sèche et inadapté à une cargaison fragile. Il n'y a pas de bon ou de mauvais navire en soi, mais un navire approprié à un voyage et à ce qu'il transporte ; il n'y a pas davantage, on le verra, de bon format en soi.

Cet article transpose la même obligation au dirigeant. C'est une obligation de moyens, et non de résultat. Le format idéal n'existant pas, il ne lui est pas demandé d'avoir choisi le bon, mais d'avoir posé la question du contenant avant que la cargaison ne parte, au moment où ce choix engage toute la suite. Dans notre métaphore, il est à la fois l'armateur et le commandant : il répond du navire comme de la cargaison. Le droit de la mer a formulé cette double responsabilité pour le contenant physique, alors que le droit du numérique n'en connaît pas l'équivalent pour le contenant numérique. L'examen des textes montrera jusqu'où il s'en approche et où il s'arrête.

1.2 Stocker n'est pas conserver

Ce qui devient illisible ne peut plus être considéré comme conservé. La formule sépare deux opérations que le langage courant confond : stocker et conserver. Stocker, c'est garder les fichiers en lieu sûr, en plusieurs exemplaires, selon les meilleurs standards techniques. Conserver est autre chose : c'est garder la capacité de leur rendre un sens. La première opération relève de l'hébergement et de la technique ; le droit s'en préoccupe beaucoup. La seconde relève au contraire du contenant, dont le droit comme les organisations se soucient peu. Un coffre dont on a perdu la combinaison protège admirablement

ce qu'il renferme ; il ne sert pourtant plus à personne et devient un coût inutile, une pièce de musée du numérique.

Par rapport à l'écrit imprimé, la différence est ici de nature. Un livre, une fois imprimé, est un objet qui demeure. Sa conservation est passive : elle ne demande guère qu'un abri, et le livre se lit encore des siècles plus tard avec les seuls yeux de son lecteur. À l'inverse, une page ou un site n'existent qu'aussi longtemps que quelqu'un les maintient : un serveur à entretenir, un nom de domaine à renouveler, un logiciel que l'on doit savoir encore faire fonctionner. Leur conservation est active et, le plus souvent, payante. Un abonnement non renouvelé ou une entreprise qui ferme suffisent à les faire disparaître.

Une étude du Pew Research Center publiée en 2024 a établi que 38 % des pages web qui existaient en 2013 étaient inaccessibles dix ans plus tard. Le chiffre mesure la disparition d'une adresse, qui n'emporte pas nécessairement celle de toute copie : des archives du web en ont parfois conservé une trace. Mais cette trace est souvent dégradée. Une page web n'est pas un objet unique : le navigateur l'assemble au moment de la lecture à partir d'images, de mises en forme et de programmes chargés séparément. Les robots d'archivage, ces programmes qui parcourent le web pour en enregistrer les pages, les capturent rarement tous et rarement au même moment. Selon une autre étude, publiée en 2015, environ 20 % seulement des pages archivées étaient à la fois complètes et fidèles à ce qu'elles avaient été. La copie existe, mais elle ne restitue pas toujours ce qui a été montré.

Certaines œuvres n'ont même plus de copie qui fonctionne. Pendant près de vingt-cinq ans, une part considérable de la création originale sur internet a été conçue pour un seul lecteur, Flash, module édité par Adobe qu'il fallait installer dans le navigateur : jeux, animations, œuvres interactives, sites entiers. Adobe a cessé d'en assurer la maintenance à la fin de 2020, puis a bloqué l'exécution de ces contenus au début de 2021. Les navigateurs, du reste, l'avaient déjà abandonné, comme ils avaient abandonné auparavant les petits programmes écrits en Java, et les archivistes avaient toujours eu du mal à capturer ces contenus avec leurs outils habituels. Dès lors, ce qui en reste ne fonctionne plus aujourd'hui que grâce à des programmes développés par des bénévoles, qui imitent l'ancien lecteur avec plus ou moins de fidélité. Le contenant a disparu et a emporté avec lui une part de la création d'une génération entière.

L'État archive d'ailleurs le web français au moyen de ces mêmes robots. Depuis 2006, la Bibliothèque nationale de France (BnF) en assure le dépôt légal, comme elle le fait depuis des siècles pour les livres. Mais le mécanisme a changé de nature. L'éditeur d'un livre dépose ses exemplaires ; le producteur d'un site n'a aucune démarche à accomplir, puisque c'est un robot qui vient photographier le web, par grandes campagnes annuelles et par collectes ciblées. Le dépôt légal du livre conserve l'objet ; celui du web conserve des instantanés, avec les lacunes que l'on vient de décrire.

L'enjeu dépasse donc les organisations, il touche l'Histoire elle-même. Depuis le tournant du siècle, les décisions des gouvernements, les débats des assemblées, les correspondances des dirigeants, la parole publique et privée ont été confiés, dans une part toujours croissante, à des contenants numériques. L'historien qui voudra écrire l'histoire de la France et du monde à partir des années 2000 disposera seulement de ce qui aura été gardé et sera encore lisible. Nul ne sait aujourd'hui quelle part en survivra. Les siècles précédents lui ont légué des archives dont la conservation demandait un abri plus qu'un entretien. Le nôtre, lui, ne lui laissera que ce que quelqu'un, à chaque génération d'outils, aura pris la peine de maintenir, de payer et de rendre encore lisible. Pour la première fois peut-être, la mémoire d'une époque dépend moins de ce qu'elle a produit que du soin qu'elle aura pris de ses contenants.

Le dirigeant est, en ce sens, un dépositaire. Il reçoit un contenu qu'il doit pouvoir rendre intact à ses successeurs et, selon le cas, à ses administrés ou à ses associés; ce qu'il rendra dépend du contenant dans lequel il l'aura gardé. L'administration le sait pour ses archives. En revanche, l'entreprise le découvre souvent plus tard, parfois trop tard. C'est le jour où une délibération, un contrat ou un élément de preuve doit être produit, des années après son établissement, devant un juge ou un repreneur. Il ne suffit plus alors de l'avoir retrouvé : encore doit-on pouvoir l'ouvrir et en établir la teneur.

Parmi toutes les qualités d'un contenant, facilité d'accès, capacité à passer d'un logiciel à l'autre, rapidité, coût, la durabilité occupe une place à part, parce qu'elle commande les autres. L'utilisateur constate chacune de ces qualités à un instant donné : aujourd'hui, ce fichier s'ouvre partout et coûte peu. La durabilité pose une autre question : ces qualités tiendront-elles ? Un format peut parfaitement circuler aujourd'hui et n'être plus ouvert par aucun logiciel en service dans vingt ans ; il était alors accessible, mais d'une accessibilité à durée limitée. Un autre peut être léger et universellement lu, tout en dépendant d'un seul acteur dont rien ne garantit la pérennité. Chacune ne vaut donc qu'à proportion du temps pendant lequel elle se maintiendra. La durabilité n'est pas une qualité de plus : elle est l'horizon à l'aune duquel toutes les autres doivent être jugées.

La qualité, au sens propre du terme, entre d'ailleurs dans ce calcul. Une photographie ou un document numérisé en haute définition occupe beaucoup de place, mais reste réutilisable : on pourra l'imprimer ou le retravailler. La même image, compressée au maximum pour économiser de l'espace de stockage, devient une vignette floue dont on ne tirera plus rien. Cette économie, des années plus tard, prive le document de ce qui faisait sa valeur. Là encore, le réglage par défaut du logiciel a souvent tranché à la place de celui qui enregistrait. Les photographes, amateurs comme professionnels, connaissent bien ce dilemme. Le format brut, dit RAW, conserve tout ce que le capteur de l'appareil a enregistré et permet de retravailler l'image des années plus tard. Mais chaque fabricant

a le sien, que seul un logiciel adapté sait lire. À l'inverse, le JPEG est lisible partout, mais il a déjà sacrifié une part de l'image au moment même où elle était prise. Ainsi choisit-on, sans toujours le savoir, entre la qualité et la compatibilité, entre la richesse de ce qu'on garde et l'indépendance à l'égard d'un fabricant.

Le cinéma et la télévision en offrent une démonstration à grande échelle. Les films tournés sur pellicule, des années 1940 aux années 1990, se restaurent aujourd'hui remarquablement. La pellicule enregistrait en effet bien plus de détails que les écrans de l'époque ne pouvaient en montrer, et il suffit de la numériser de nouveau pour en tirer une image de haute définition. Il n'en va pas de même des productions terminées sur bande vidéo, à la définition des téléviseurs de leur temps. L'exemple le plus connu est celui d'une célèbre série de science-fiction américaine. Sa première déclinaison des années 1980 et 1990 a pu être entièrement reconstruite en haute définition, au prix d'un travail considérable, parce que ses images et ses maquettes existaient encore sur pellicule. En revanche, ses deux suites, dont les effets spéciaux avaient été calculés par ordinateur et intégrés directement sur bande vidéo, ne peuvent pas l'être sans refaire intégralement ces effets. La technique la plus moderne de l'époque, choisie parce qu'elle était plus rapide et moins chère, a vieilli plus mal que le support ancien qu'elle remplaçait. L'économie du jour a été payée en longévité.

Cet horizon n'impose pas pour autant un contenant unique. Il serait absurde d'exiger d'une note de service qu'elle soit conservée comme un acte notarié ou d'un tableau de travail qu'il soit figé comme une pièce de preuve. Le document vivant demande à être modifié; la pièce de preuve, à ne plus l'être; l'archive, à traverser le temps sans altération. Dès lors, la durabilité oblige à examiner, pour chaque cargaison, la traversée qu'elle doit accomplir. On retrouve la navigabilité relative au voyage : le dirigeant choisit le contenant en fonction de la route; la route la plus longue est celle du temps. Un armateur ne vérifie pas seulement que son navire flotte au port; il s'assure qu'il tiendra la traversée.

Cet horizon explique aussi pourquoi l'hébergement, si soigneusement encadré, ne suffit pas. Un ensemble de documents peut être hébergé sur une infrastructure irréprochable, dans un pays maîtrisé, par un prestataire qualifié. Il deviendra pourtant illisible si le contenant qui le porte cesse d'être lu; il deviendra inutilisable si la clef qui permet de l'interroger vient à manquer. Le meilleur des coffres ne protège pas d'une langue oubliée.

Cette doctrine prolonge celle que la revue a posée dans son premier hors-série, *La sûreté de l'intention*. Elle en est le complément : l'intention d'un dirigeant ne reste fidèle que si le contenant qui la porte demeure lisible. Une décision exactement formulée, mais consignée dans un contenant qui deviendra illisible, n'aura été conservée que le temps que ce contenant aura duré.

1.3 Le choix du réglage par défaut

Ce choix du contenant, dont tout dépend, n'est presque jamais exercé par celui qui en répond. Un choix non exercé ne cesse pas pour autant d'exister : il est exercé par quelqu'un d'autre, ailleurs, selon d'autres critères. Ne pas décider est encore une manière de décider, et subir les réglages par défaut revient à les choisir sans le savoir.

Le geste le plus fréquent de la bureautique en offre la démonstration la plus simple. Celui qui clique sur « Enregistrer » accepte un format que l'éditeur de son logiciel a fixé des années plus tôt, pour des millions d'utilisateurs, sans rien savoir du document en cause. Celui qui exporte un tableau accepte l'encodage que l'outil lui propose. L'utilisateur d'un espace de travail partagé hérite des réglages retenus par l'administrateur. Quant à celui qui puise dans un ensemble de données livré « prêt à l'emploi », il hérite de tous les choix faits en amont par celui qui l'a préparé. À aucun de ces moments le dirigeant n'a décidé, et à chacun d'eux une décision a été prise.

Quatre acteurs se partagent ainsi un choix qu'aucun d'eux n'a entendu faire. L'éditeur du logiciel décide par son réglage par défaut ; l'administrateur informatique, par sa configuration ; la plateforme, par ce qu'elle livre prêt à l'emploi ; le technicien, enfin, par ses habitudes de travail. Quand le résultat déçoit, l'organisation met pourtant en cause le technicien et se trompe de cible.

Le technicien ne trahit personne. Il optimise ce sur quoi sa hiérarchie le juge : le respect du délai, la compatibilité avec les outils déjà en place, l'absence de plainte des utilisateurs, la facilité de maintenance. Ces critères sont légitimes ; ce sont d'ailleurs les seuls dont il dispose lorsqu'aucun autre ne lui a été donné. Faute de consigne sur la durabilité, la réversibilité ou la dépendance, il retient la solution la plus simple selon ses propres critères. Or ceux-ci ne coïncident pas nécessairement avec l'intérêt de l'organisation qu'il sert. Certes, son choix est rationnel ; mais il n'est pas celui qu'aurait fait le dirigeant s'il avait pris la peine de choisir.

Les économistes ont donné un nom à ce mécanisme : la théorie de l'agence. Deux économistes américains, Michael Jensen et William Meckling, l'ont formalisée en 1976. Ils y décrivent la relation par laquelle une personne (le mandant) confie à une autre (l'agent) le soin d'agir en son nom, en lui déléguant une part de son pouvoir de décision. Le mandant est ce que la littérature économique appelle le principal. Ils en concluaient que, chacun poursuivant son propre intérêt, l'agent n'agirait pas toujours au mieux de celui du mandant. Or le cas du contenant en présente une variante. L'agent n'y poursuit aucun intérêt contraire à celui du mandant. Il applique seulement des critères différents, les siens, parce que le mandant ne lui en a pas donné d'autres. La divergence ne naît pas d'une opposition d'intérêts ; elle naît du silence du mandant.

La responsabilité remonte alors au dirigeant. Elle y remonte parce que, sans avoir rien choisi, il n'a pas non plus donné de consigne. Le défaut tient à l'absence du cap qui aurait permis au technicien de décider selon les critères de l'organisation. La revue décrivait déjà ce ressort dans son deuxième numéro, à l'échelle des organes de gouvernance : un choix que personne ne conteste parce que chacun suppose qu'un autre l'a examiné. Le contenant en est la déclinaison à l'échelle technique. Le dirigeant suppose que le technicien a pesé la durabilité, le technicien suppose que la question ne lui était pas posée, et chacun a raison de son point de vue.

On distingue d'ordinaire deux figures : le réglage par défaut et le prêt à l'emploi. Le premier paraît anodin : un paramètre proposé à l'installation d'un logiciel. Le second paraît un service : un ensemble de données ou un outil livré préparé, qui épargne à son utilisateur un travail fastidieux. Les deux relèvent pourtant d'un seul et même geste, source d'une même dépendance cachée : un arbitrage rendu une fois, en amont, sans délibération, et imposé à tous ceux qui bâtissent dessus. Ils ne diffèrent que par l'échelle. Le réglage par défaut décide du contenant d'un document ; le prêt à l'emploi décide du contenant de tout ce qui sera construit à partir de lui. On en verra plus loin la forme la plus contemporaine, au cœur même de l'État.

Qu'on se représente un armateur qui confierait son navire à un équipage sans lui indiquer ni la destination, ni la nature de la cargaison, ni les ports à éviter. L'équipage ne resterait pas à quai pour autant. Il suivrait la route la plus habituelle, la plus facile, celle que les cartes de bord proposent d'emblée et que suit quiconque n'en a pas choisi d'autre. Il arriverait peut-être à bon port. Mais ce port ne serait pas nécessairement celui que l'armateur aurait choisi. Et l'armateur ne pourrait reprocher à l'équipage d'avoir mal exécuté des ordres qu'il ne lui avait jamais donnés.

On dit volontiers que la facilité a pour prix l'enfermement dans un format propriétaire. C'est inexact : le chemin le plus aisé peut conduire à un format ouvert, et le web tout entier, bâti sur un format ouvert, en est la preuve vivante. La facilité a un autre prix, plus certain encore : une délégation. Car le chemin facile est celui que quelqu'un d'autre a déjà choisi. L'économie numérique en offre de nombreux exemples : les extensions payantes et les services « clés en main » qui habillent, pour le rendre plus commode, un logiciel libre disponible gratuitement par ailleurs. On paie l'habillage et la commodité, et, avec eux, la dépendance envers celui qui les fournit. En empruntant le chemin le plus facile, on perd moins une liberté technique qu'une décision. À l'inverse, l'indépendance coûte l'effort de reprendre chacun de ces choix, étage par étage, et d'en chiffrer la sortie.

Ce constat ne condamne pas pour autant la délégation, sans laquelle aucune organisation ne fonctionnerait. Nul dirigeant ne peut arrêter lui-même le format de chaque fichier ni le paramétrage de chaque outil, et il serait déraisonnable de le lui demander. Ce qui

lui est demandé est autre : que le choix du contenant soit délégué, et non abandonné. La différence tient en un mot, le mandat. Une délégation fixe les critères de celui qui décide ; un abandon le laisse décider selon les siens. Le geste propre du dirigeant, en matière de contenant, est donc de fixer le cap ; la conclusion y reviendra.

1.4 Sous la ligne de flottaison

Pourquoi ce cap fait-il si souvent défaut ? Pour une part, c'est que rien ne le réclame. Le dirigeant attentif à ses obligations se tourne vers la norme pour savoir ce qu'on attend de lui. Or la norme, en matière numérique, lui parle avec insistance d'un seul étage et presque pas des autres.

Trois regards doivent ici être distingués. Le décideur voit le navire : le logiciel qu'il utilise, l'écran qu'il a sous les yeux. Le droit voit le pavillon et le port d'attache : le lieu où les données sont hébergées et la loi à laquelle elles sont soumises, que le décideur lui-même ne voit guère. Presque personne ne voit sous la ligne de flottaison, là où se jouent le format des fichiers, le modèle d'intelligence artificielle qui les classe et la plateforme par laquelle ils circulent. On les décrira un à un.

Cette répartition entre trois regards explique un paradoxe apparent. L'hébergement est, matériellement, l'étage le plus profond de toute la chaîne : des serveurs enfermés dans des centres de données, qu'aucun utilisateur ne voit jamais. Il est pourtant, pour le droit, l'étage le plus visible, et pour une raison simple : c'est aussi le plus matériel. Des machines, des bâtiments, un territoire, un exploitant soumis à une loi : le droit sait saisir une chose située quelque part, comme l'administration maritime inscrit un navire au registre de son pavillon et l'inspecte au port. Ses enjeux, la localisation des données et leur exposition à des lois étrangères, relevaient des catégories familières du droit public : la souveraineté et la compétence territoriale des États. En revanche, ceux du format et du modèle, qui n'ont pas de lieu, y entraient mal et sont restés dans l'ombre.

De là procède ce que cet article appellera la « loi du visible » : la norme règle ce qu'elle voit. Le constat n'impute aucune intention à personne. Il décrit seulement un champ de vision. Il en résulte une erreur d'objet : la norme saisit le maillon de la chaîne qui lui est visible, et non celui où se trouve le levier qui détermine réellement ce que le lecteur pourra lire et retrouver. Le mécanisme n'est pas propre au numérique ; la revue le retrouvera sur d'autres terrains. Le numérique en offre toutefois l'une des démonstrations les plus nettes.

La contrainte que la norme exerce faiblit alors dans deux directions. Elle décroît à mesure qu'elle approche du réglage par défaut qu'elle prétend gouverner : affirmée dans ses principes, elle n'atteint presque jamais le paramètre effectivement appliqué. Elle décroît aussi à mesure qu'on s'éloigne de l'hébergement. La règle est obligatoire pour le pavillon,

l'hébergement : la loi de 2024 impose aux administrations de protéger leurs données sensibles contre les lois étrangères. De même, le règlement général sur la protection des données (RGPD) encadre depuis 2018 les transferts de données personnelles hors de l'Union européenne. Elle est simplement incitative pour la coque, le format des fichiers : la loi de 2016 se borne à « encourager » les formats ouverts. Elle est muette pour la carte, le modèle qui permet de retrouver les documents, comme sur le choix du détroit, c'est-à-dire de la plateforme par laquelle ils circulent. Le paradoxe central de cet article naît de cette double faiblesse. Une administration peut respecter entièrement l'obligation d'héberger ses données en France, sous la protection la plus exigeante. Elle peut pourtant confier à un modèle d'intelligence artificielle conçu à l'étranger le soin de décider quels textes de loi s'afficheront lorsqu'un agent interroge le droit en vigueur.

2 Les étages de la chaîne numérique

Avant d'examiner ce que le droit en voit, la chaîne doit être décrite en entier. Suivre une donnée de sa production à sa lecture, c'est suivre une cargaison de l'embarquement jusqu'au quai de déchargement. Sept étages se succèdent sur ce trajet. Chaque étage a sa fonction et installe une dépendance qui lui est propre ; chacune de ces dépendances peut être choisie ou subie. Pour chaque étage, on exposera sa nature, sa place dans la métaphore, un fait observé qui l'éprouve et la dépendance qu'il crée.

La métaphore distribue ainsi les rôles, et chacun d'eux correspond à une réalité précise. La donnée est la cargaison : les chiffres, les textes ou les images que l'on veut transmettre. Le format du fichier est la coque qui la porte : la manière dont ces données sont inscrites et organisées dans le fichier. L'outil est le chantier naval qui a construit cette coque : le logiciel avec lequel le fichier a été produit, et qui en a fixé le format. Le modèle d'enrichissement est la carte marine grâce à laquelle on retrouve son chemin dans la cargaison : un programme d'intelligence artificielle qui classe les textes selon leur sens pour permettre d'y chercher. Il ne se confond pas avec les métadonnées, ces informations descriptives (auteur, date, titre) que l'on attache à un fichier : les métadonnées décrivent un document, alors que le modèle le situe par rapport à tous les autres.

L'hébergement est le port d'attache et le pavillon sous lequel le navire est immatriculé. Il désigne la machine où le fichier réside au départ, qu'il s'agisse d'un ordinateur, d'un site internet ou d'un serveur, et le droit du pays dont relève son exploitant. Le canal de diffusion est fait des routes, des détroits et des câbles par lesquels le navire passe : à la fois les réseaux physiques de l'internet et les plateformes par lesquelles le fichier est mis à disposition. Le poste de consommation est le quai de destination, où la cargaison doit être déchargée et lue : l'appareil de celui qui reçoit le fichier, ordinateur, téléphone ou serveur, et le logiciel avec lequel il l'ouvre. La licence, enfin, n'est pas une étape du voyage. Elle est le connaissance, ce document qui, dans le transport maritime, décrit la marchandise, accompagne la cargaison de bout en bout et dit ce que chacun a le droit d'en faire. Le dirigeant est l'armateur et le commandant, et répond de tout le trajet. Or le droit, on le verra, ne regarde guère que le pavillon.

2.1 La fausse nudité de la donnée

La donnée est ce qui est produit. Elle peut être à peu près n'importe quoi : une liste de subventions, un texte de loi, un registre, une série de mesures, des données techniques ou financières, une photographie d'art en haute définition, un film. C'est la cargaison, ce qu'on veut faire arriver intact et lisible. C'est aussi le seul étage que le dirigeant surveille spontanément, parce qu'il se confond avec le contenu dont il doit répondre.

Pourtant, la donnée ne naît jamais nue. Un montant saisi avec la virgule décimale française devient, pour un logiciel réglé selon la convention anglo-saxonne, un simple texte, voire deux nombres distincts. Une date écrite en toutes lettres n'est pas une date pour la machine qui voudra la trier. Dès sa saisie, la cargaison est emballée d'une certaine manière, et cet emballage décide en partie de la façon dont elle supportera le voyage.

Deux exemples en donnent la mesure. Les codes des communes françaises commencent souvent par un zéro : celui de la première commune de l'Ain s'écrit 01001. Un tableur qui croit reconnaître un nombre supprime ce zéro, et le code devient 1001, qui ne désigne plus rien. Le phénomène a pris des proportions spectaculaires en génétique. Plusieurs gènes humains portaient des noms que les tableurs transformaient d'office en dates, comme SEPT1 ou MARCH1. Une étude publiée en 2016 dans la revue *Genome Biology* a constaté qu'environ un cinquième des articles scientifiques accompagnés de listes de gènes sous forme de tableur contenaient de telles erreurs. En 2020, l'organisme international chargé de nommer les gènes humains a fini par en rebaptiser plusieurs : il était plus simple de changer le nom des gènes que le comportement des logiciels.

2.2 Le format de la traversée

Le fichier est l'enveloppe du contenu : il en est le contenant. Mais un fichier n'est, pour la machine, qu'une suite de nombres. C'est le format du fichier qui dit comment les lire et quoi en faire, un texte, un tableau ou une image, comme une partition n'est qu'une suite de signes pour qui ignore le solfège. L'extension qui termine le nom du fichier signale d'ordinaire le format. Ces quelques lettres, que chacun a sous les yeux sans y prêter attention, sont .doc ou .docx pour un document Word, .odt pour son équivalent libre, .xls et .xlsx pour un tableur, .ppt et .pptx pour une présentation. On connaît aussi .pdf, .txt, .csv, .html pour une page web, .jpg ou .png pour une image. Dans notre métaphore, le format du fichier est la coque, avec son tirant d'eau et sa taille. Invisible sous la ligne de flottaison, elle décide pourtant des ports où l'on peut entrer et de ce qu'on peut porter.

Toutes les extensions ne désignent d'ailleurs pas un format de document. Un fichier .zip ou .rar n'est pas un document mais une caisse : une archive compressée qui en contient d'autres, chacun dans son propre format. À l'inverse, certains documents n'ont pas de fichier propre tant qu'ils ne quittent pas leur logiciel. Un texte rédigé dans Google Docs vit sur les serveurs de Google et ne devient un fichier qu'à l'exportation, dans un format qu'il faut alors choisir. Le traitement de texte d'Apple produit des fichiers .pages que les logiciels des autres éditeurs lisent mal ou pas du tout.

Tout en bas de cet étage se trouve l'encodage des caractères, cette table de correspondance entre nombres et lettres évoquée en introduction. Deux tables différentes lisent les mêmes nombres comme deux textes distincts : un « é » devient ainsi, à l'arrivée, « Ã© »

ou un point d'interrogation.

Or le défaut ne vient pas toujours de celui qui écrit. Un fichier du répertoire national des élus, téléchargé le 20 septembre 2026 sur la plateforme nationale, est écrit dans l'encodage universel, l'UTF-8, conformément à la règle. Ouvert dans un tableur réglé selon l'ancienne table de Windows, il affiche pourtant « dÃ@partement » au lieu de « département » dès l'en-tête. La raison en est minuscule : le fichier ne porte pas la petite marque facultative qui, placée en tête, annonce au logiciel l'encodage employé, et le logiciel applique alors son propre réglage par défaut. Ainsi le producteur a respecté la norme ; le lecteur voit pourtant un texte abîmé. La conformité ne suffit donc pas : le contenant doit aussi être jugé à l'autre bout de la chaîne, sur le quai, là où un réglage que le lecteur n'a pas choisi décide de ce qui s'affiche.

Le SMS en offre une illustration que chacun a dans sa poche. Il repose sur une conception de la fin des années 1980, qui a fixé une fois pour toutes la taille d'un message : 140 octets, soit 160 caractères dans l'alphabet réduit d'origine. Un SMS long d'aujourd'hui n'est en réalité qu'une succession de petits SMS, que le téléphone du destinataire réassemble en un seul message grâce à une étiquette glissée dans chaque morceau. Chaque morceau ne contient plus alors que 153 caractères. Le fractionnement a cessé d'être visible pour l'utilisateur ; il n'a pas cessé d'exister, et les plateformes professionnelles d'envoi facturent toujours chaque morceau comme un SMS entier. En outre, un seul caractère étranger à l'alphabet d'origine suffit à faire basculer tout le message dans un encodage plus riche, qui ne loge plus que 70 caractères par SMS et 67 par morceau. Ce peut être un émoji, un tiret long ou cette apostrophe courbe que de nombreux claviers insèrent d'eux-mêmes à la place de l'apostrophe droite. Un réglage de clavier que l'utilisateur n'a pas choisi divise ainsi par plus de deux la capacité d'un contenant conçu il y a près de quarante ans et devenu invisible. Le contenant survit à son époque et gouverne encore ce qu'il transporte.

L'étage du format a été beaucoup débattu, en l'associant au mot « ouverture » : un format ouvert serait bon, un format fermé mauvais. Le mot est trompeur, parce qu'il recouvre sept questions distinctes que le débat public confond. La première est celle de la gouvernance : qui décide de l'évolution du format, un éditeur seul ou une instance où plusieurs parties délibèrent ? La deuxième est celle de l'accès : sa description technique est-elle publiée, complète, consultable sans frais ? La troisième est celle de la dépendance d'outil : peut-on produire et lire ce format avec plusieurs logiciels indépendants les uns des autres ? La quatrième est celle de la destination : le format est-il fait pour être modifié, ou pour figer une présentation ? La cinquième est celle de la pérennité : sera-t-il encore lisible dans vingt ans ? La sixième est celle de l'interopérabilité réelle, qui n'est pas l'interopérabilité promise : les échanges entre logiciels se font-ils sans perte ? La septième, enfin, est celle du coût de sortie : que coûterait un changement ?

Or ces sept questions ne reçoivent pas la même réponse pour un même format. Un format peut être irréprochable sur l'une et fragile sur l'autre. C'est pourquoi la question « ce format est-il ouvert ? », posée en ces termes, n'appelle aucune réponse utile. Mieux vaut demander : ouvert en quoi, et pour quoi faire ? La ligne de partage décisive n'est pas l'ouverture formelle, qu'une norme internationale suffit à conférer. C'est la gouvernance réelle : qui, dans les faits, commande l'évolution du format et le réglage par lequel il est produit ?

Le principe posé plus haut, qu'il n'existe pas de bon navire en soi, vaut pour les formats, et l'image d'une flotte permet de les ranger par usage plutôt que par mérite.

Le *cargo polyvalent*, ce sont les formats bureautiques : ceux des traitements de texte, des tableurs et des présentations. Ils portent de tout, mais leur coque dépend du chantier qui les construit. Deux familles se partagent la mer. La première, le format OpenDocument (fichiers .odt, .ods et .odp), évolue au sein d'un consortium où siègent de nombreux acteurs. La seconde, le format Office Open XML (fichiers .docx, .xlsx et .pptx), est celle de Microsoft Office. Toutes deux ont reçu le label d'une norme internationale. Mais la seconde ne dissimule pas sa généalogie : son texte même lui assigne pour objet de représenter fidèlement l'ensemble des documents produits par les anciennes versions du logiciel de son éditeur. Elle existe pour cela en deux variantes, dont l'une, dite transitoire, conserve les particularités de ces anciens formats. Rien de cela n'est irrégulier. Mais une norme dont l'objet est la continuité d'un ensemble existant de documents reste, dans les faits, ordonnée aux besoins de celui qui a produit cet ensemble.

Le détail le plus parlant tient au réglage par défaut. Les logiciels produisent ordinairement la variante transitoire plutôt que la variante stricte. Des tests conduits en 2017 par la Bibliothèque du Congrès américaine ont montré que ce choix n'était pas propre à un seul éditeur. Google Docs et LibreOffice créaient eux aussi leurs nouveaux documents dans la variante transitoire, comme Word le fait par défaut depuis 2013. Lorsque Word propose de convertir un ancien document au format le plus récent, c'est encore cette variante transitoire qu'il produit. Les deux variantes partagent pourtant la même extension, qui ne permet pas de les distinguer. Dès lors, celui qui reçoit un fichier .docx ne sait pas, à son nom, lequel des deux navires il accueille. Quant aux anciens formats, .doc, .xls ou .ppt, que l'on croyait disparus, ils circulent encore : on en trouve par dizaines parmi les documents publiés aujourd'hui par les services de l'État.

Le *cargo au long cours*, c'est le PDF, et surtout sa version conçue pour l'archivage, le PDF/A. Il est fait pour figer une présentation, non pour être modifié : on peut y retoucher un mot, mais mal, et au risque de défaire la mise en page. Taillé pour traverser le temps sans altération, il embarque ce qui lui est nécessaire pour s'afficher à l'identique, jusqu'aux polices de caractères : la norme d'archivage impose d'y intégrer les polices employées,

le plus souvent réduites aux seuls caractères utilisés. Sa force est la fixité, sa faiblesse la modification. C'est le navire de la pièce de preuve et de l'archive; le document vivant en demande un autre. Pour l'image, le format TIFF joue un rôle comparable : standard de la numérisation dans les années 1990, il reste lisible aujourd'hui et demeure employé par les services d'archives.

La *vedette*, c'est le fichier texte brut (.txt) et son cousin le Markdown : une manière d'écrire un texte structuré, titres et listes compris, que l'œil lit sans aucun logiciel. Légère et rapide, elle passe partout, mais elle ne porte qu'une cargaison peu structurée. Le Markdown n'a pas de maître, ce qui fait sa liberté. Il la paie en compatibilité incertaine, des variantes étant apparues au point qu'une description commune a dû être entreprise après coup. Ce type de navire enseigne qu'il y a deux manières d'être libre, par une gouvernance partagée ou par l'absence de gouvernance, et qu'elles n'offrent pas les mêmes garanties.

La *péniche*, c'est le CSV, ce fichier « à plat » où les valeurs d'un tableau sont écrites en simples lignes de texte, séparées par un caractère convenu, et que tous les tableurs savent produire et lire. Son nom annonce la virgule (*comma-separated values*); en pratique, le séparateur est tantôt la virgule, tantôt le point-virgule, que les tableurs français emploient parce que la virgule y sépare les décimales, tantôt la tabulation. Le CSV a ainsi plusieurs identités, et rien dans le fichier ne dit laquelle il a adoptée. Simple et économique, il est le format par excellence de la donnée publique. Mais il ne navigue que sur les voies intérieures et s'échoue dès que le tirant d'eau de son encodage ne correspond pas à la profondeur de la voie, autrement dit au réglage du logiciel qui l'ouvre. Le CSV ne possède pas de norme qui fasse vraiment autorité, ni de signature, ces premiers octets par lesquels un fichier annonce ce qu'il est. Surtout, rien en lui ne dit dans quel encodage il a été écrit. Il ressemble à une lettre qui ne dirait pas dans quelle langue elle est rédigée : le destinataire doit deviner et, s'il se trompe, lit des signes dépourvus de sens sans que rien ne l'avertisse de son erreur.

La péniche est aussi le navire des fausses déclarations les plus banales. Une ressource publiée par la ville de Béthune se présente comme un fichier Excel classique, alors qu'elle est écrite dans un ancien format du même éditeur, depuis remplacé, qui voyage ici sous l'étiquette d'un autre. Le fait n'a rien de grave et ne vise personne; il est parlant par sa banalité même. L'extension d'un fichier n'est qu'une déclaration de marchandise, qu'aucun contrôle ne vérifie à l'embarquement. Chacun peut en faire l'expérience : renommer un fichier .doc en .xls, l'ordinateur l'accepte sans objection; c'est à l'ouverture que le tableur échoue.

Les images ont leur propre flotte, familière sans qu'on y pense. Les extensions .jpg et .jpeg désignent le même format, la plus courte étant héritée des systèmes qui limitaient l'extension à trois lettres. Le .jpg compresse l'image en sacrifiant une part de ses détails,

d'autant plus que le réglage choisi est économe ; le .png la conserve sans perte, au prix d'un fichier plus lourd. Depuis quelques années, les iPhone enregistrent par défaut leurs photographies dans un troisième format, plus compact, que certains ordinateurs ne savent pas ouvrir sans conversion. Certes, aucun de ces choix n'est mauvais en soi. Chacun engage pourtant la durée de vie et la réutilisation de l'image, et presque aucun n'est fait par celui qui appuie sur le déclencheur.

La flotte compte enfin des *voiliers*. Le vent n'appartient à personne, et un voilier, du dériveur d'école au cargo à voile, peut aller partout sans dépendre d'aucun fournisseur de carburant. Leurs équivalents numériques sont les formats dont la règle est publique et libre de droits, que n'importe quel logiciel peut lire et écrire. Le HTML, le format des pages web, est ouvert et porte le web tout entier. Le JSON, format d'échange de données entre logiciels, impose par sa norme même l'encodage universel pour tout échange entre des systèmes qui n'appartiennent pas à un même ensemble clos. Il règle ainsi, une fois pour toutes, la question que le CSV laisse ouverte à chaque fichier.

Reste la taille, que l'image du tonnage rend sensible. Un cargo trop lourd ne remonte pas le fleuve, et un fichier trop lourd n'est pas lu par un outil léger ; il exige aussi des infrastructures à sa mesure, pour être stocké, transmis et reçu. La taille commande aussi la vitesse. Lorsque des logiciels échangent des données en continu, par ces interfaces qui permettent à un programme d'interroger une base, un contenant trop lourd ralentit chaque échange, jusqu'à rendre le service inutilisable. Ainsi le choix du contenant dépend-il aussi de la vitesse dont on a besoin. On verra plus loin que certains ensembles de données publiques sont désormais livrés avec, pour chaque passage de texte, une longue suite de nombres produite par une intelligence artificielle pour permettre la recherche par le sens. Cette charge alourdit chaque fichier pour tous ceux qui le téléchargent, qu'ils en aient besoin ou non.

La leçon de la flotte est celle d'une navigabilité relative. La question que le dirigeant doit poser n'est jamais « quel est le meilleur format ? ». Elle est : « pour quelle traversée, et avec quelle cargaison ? »

2.3 Le gabarit du logiciel

L'outil est le logiciel qui produit le fichier, du côté de l'émetteur. Celui qui l'ouvre à l'arrivée peut être différent : un PDF produit avec Word se lit dans n'importe quel lecteur de PDF. Tout se décide pourtant du côté de l'émetteur, et c'est ce côté qu'examine le présent étage ; celui du destinataire relève du poste de consommation. Dans notre métaphore maritime, il est le chantier naval ; or un chantier construit par défaut à son propre gabarit, selon ses propres normes, sauf si on l'oblige à faire autrement. Le logiciel impose son format comme un chantier impose les dimensions de ses cales, sans que l'armateur ait

eu à le demander. C'est à cet étage que l'éditeur fabrique le réglage par défaut. Trois cas, choisis parce qu'ils semblent se contredire, en montrent la mécanique.

Le premier est celui d'Euro-Office, suite bureautique lancée en 2026 par une coalition d'acteurs européens pour proposer une alternative aux outils de Microsoft. Elle est construite à partir du code d'un logiciel libre existant ; un logiciel libre est un logiciel dont chacun peut lire, modifier et redistribuer le code. Pour permettre aux organisations de quitter l'éditeur dominant, elle doit d'abord parler sa langue ; elle privilégie donc, dans son réglage par défaut, le format de cet éditeur, celui de la suite Microsoft. Ce n'est pas une erreur de conception, c'est la condition même de son adoption : une alternative qui n'ouvrirait pas les fichiers existants, créés sous Microsoft, ne serait adoptée par personne. Mais la compatibilité qui rend la sortie possible entretient aussi le format dont on voulait sortir. Et dans la plateforme collaborative où elle s'installe, le choix entre elle et une autre suite libre est renvoyé à l'administrateur informatique de chaque installation : aucune autorité ne le tranche pour l'ensemble. Le choix abandonné trouve ici sa forme exacte.

Le deuxième cas est celui de Grist, un tableur collaboratif conçu par un éditeur new-yorkais et adopté par l'État français, qui l'héberge lui-même et contribue à son développement. Son code est ouvert : chacun peut le lire, le modifier et le reprendre. Si l'éditeur venait à changer de cap, l'administration pourrait poursuivre l'outil sans lui, en reprenant le code : c'est ce que les développeurs appellent une bifurcation. L'outil est étranger, mais la dépendance est réversible.

Le troisième cas apporte la confirmation de l'histoire. La bibliothèque cartographique MapLibre, qui sert à afficher des cartes sur le web, est née d'une telle bifurcation, le jour où l'outil dont elle est issue a abandonné sa licence libre pour une licence propriétaire. L'Institut national de l'information géographique et forestière (IGN) l'a retenue pour son portail cartographique national. La reprise n'a été possible que parce que l'ancienne licence le permettait : ceux qui avaient bâti sur l'outil n'ont pas eu à renégocier leur liberté, elle était inscrite dans le connaissance d'origine.

Mis côte à côte, ces trois cas défont une intuition répandue. Une suite européenne peut reconduire un verrou, un outil américain peut laisser la porte ouverte. Dès lors, la nationalité de l'outil est un indicateur, elle n'est pas un critère. Le critère est la réversibilité, et une question suffit à l'éprouver : si l'on devait partir demain, que faudrait-il reprendre, en combien de temps et à quel coût ?

2.4 La carte du sens

L'étage du modèle d'enrichissement est le plus récent et le plus méconnu. Pour chercher dans un grand ensemble de textes par le sens plutôt que par les mots exacts, on le fait lire

par un modèle d'intelligence artificielle, programme entraîné sur de très grandes quantités de textes. Celui-ci traduit chaque passage en une longue suite de nombres, que les spécialistes appellent un vecteur ou, en anglais, un *embedding*. Ces nombres fonctionnent comme des coordonnées : elles placent chaque passage dans un espace où les textes de sens voisin se retrouvent proches les uns des autres. Pour chercher, on traduit la question avec le même modèle et on retient les passages dont les coordonnées sont les plus proches des siennes.

La meilleure comparaison est celle d'une bibliothèque où les livres ne seraient plus rangés par auteur ou par titre, mais par proximité de sujet, selon un classement établi par un bibliothécaire dont on n'a pas choisi la méthode. Tant que l'on cherche dans cette bibliothèque, on doit passer par son catalogue et suivre sa logique. Dans notre métaphore, le modèle est la carte marine et, naturellement, celui qui dresse la carte décide de ce que l'on trouve et de ce que l'on ne trouve pas.

L'expérience en est quotidienne. Nous nous en remettons chaque jour, pour nous orienter, à un système de guidage par satellite que nous ne maîtrisons pas. Jusqu'en 2000, les États-Unis dégradaient volontairement, pour des raisons de sécurité nationale, la précision du signal civil de leur système, qui ne permettait alors de se situer qu'à une centaine de mètres près. L'Europe a depuis construit le sien, Galileo, qui fournit ses premiers services depuis 2016. Qu'on imagine un instant le guidage par satellite d'aujourd'hui redevenu inutilisable, parce qu'il ne situerait plus qu'à cent ou deux cents mètres près. Une bonne part de nos déplacements, de notre logistique et de nos services d'urgence perdrait le nord. Celui qui tient la carte maîtrise la route.

La dépendance née du modèle est d'une nature particulière, et d'autant plus insidieuse qu'elle se présente comme un service. Le modèle ne cache rien : il propose sa propre carte, et c'est à travers elle que l'on voit tout le fonds documentaire. Les coordonnées produites par un modèle ne se lisent qu'avec ce modèle, puisque la question doit être traduite dans le même système que les textes. On ne les convertit pas dans le système d'un autre modèle sans perte : il existe bien des outils de traduction d'un système à l'autre, mais ils sont approximatifs et dégradent la qualité de la recherche. Dès lors, changer de modèle impose, en pratique, de tout recalculer. La sortie est possible, mais elle a un coût ; ce coût n'est presque jamais chiffré. Ainsi naît ce que cet article appelle la « clef sémantique » : les textes appartiennent à celui qui les publie, mais la manière d'y trouver quelque chose appartient au modèle. La revue a décrit, dans *La sûreté de l'intention*, les mécanismes par lesquels un système d'intelligence artificielle peut orienter le jugement de celui qui s'en sert ; la clef sémantique en est la version documentaire.

En outre, celui qui fixe l'espace dans lequel un ensemble de textes est cherché fixe en partie ce qui sera trouvé. Les moteurs de recherche en offrent la démonstration la plus

familière. Un moteur ne montre pas le web : il en montre un classement. Or presque personne ne va au-delà de la première page : selon une étude de la société Backlinko portant sur environ 4 millions de résultats de Google, 0,63 % seulement des internautes cliquent sur un résultat de la deuxième page. Un site qui n'y figure pas existe bel et bien ; mais pour l'immense majorité de ceux qui cherchent, c'est comme s'il n'existait pas. À l'inverse, la première place ne dit rien de la qualité de ce qu'elle désigne : elle peut s'acheter, les premiers résultats étant souvent des annonces payées, ou se conquérir par un référencement habile. Le classement ne juge pas le fond ; il décide pourtant de ce que le lecteur verra. Un modèle entraîné sur d'autres langues et d'autres textes, selon d'autres priorités, rapprochera les documents selon ses propres affinités, qui ne sont pas nécessairement celles du droit français ni de ses lecteurs. Cet article nomme cela capture cognitive, sans prêter à quiconque la moindre intention : la carte n'a pas besoin d'être falsifiée pour orienter la navigation, il suffit qu'elle ait été dressée ailleurs. Le numéro précédent de la revue a décrit, sous le nom de souveraineté sémiotique emprunté à l'historien Thomas Mullaney, la maîtrise des systèmes par lesquels une langue s'écrit et se code ; la carte du sens en est le prolongement.

Depuis trois siècles, l'État tient pourtant à dresser lui-même ses cartes marines. Le Service hydrographique et océanographique de la Marine (SHOM), héritier du dépôt des cartes et plans créé sous l'Ancien Régime, exerce toujours cette mission et diffuse les informations officielles nécessaires à la navigation. Mais ce que l'on sait depuis si longtemps de la carte physique n'a pas encore été transposé à la « carte du sens », ce classement des textes par proximité de signification que dresse un modèle. On verra ce qu'il en est pour le droit français lui-même lorsqu'on examinera la manière dont l'État organise ses propres outils d'intelligence artificielle.

2.5 La loi de l'hébergeur

L'hébergement, ce sont les serveurs sur lesquels résident les données et le droit dont relève celui qui les exploite. Héberger ses données chez un prestataire, c'est un peu comme louer un entrepôt : on y dépose ses biens, mais le bâtiment, les clefs et le règlement intérieur appartiennent à un autre. Dans notre métaphore maritime, l'hébergement est le port d'attache et le pavillon, le seul étage inscrit au registre et contrôlé à quai. C'est le plus visible pour le droit et le moins visible pour l'utilisateur, qui ne sait presque jamais sur quelle machine ni sous quelle loi ses fichiers sont gardés.

La question que pose cet étage n'est pas d'abord technique, elle est juridique. Certains États se sont dotés de lois qui permettent à leurs autorités d'exiger d'une entreprise soumise à leur droit la communication de données qu'elle détient, y compris lorsque ces données sont stockées à l'étranger. Ainsi un serveur installé en France mais exploité par

une société soumise à une telle loi porte-t-il, en quelque sorte, deux pavillons. C'est cette exposition que le droit français a entrepris de traiter; on verra avec quelle vigueur.

L'hébergement pose aussi une question de durée, que l'on mesure moins. Changer d'hébergeur n'est pas un déménagement de meubles. En effet, la migration impose de transférer des volumes considérables de données, de réécrire les applications qui dépendaient des services propres au premier prestataire, de former les équipes, de tester, puis de basculer sans interrompre le service. Plus l'organisation s'est appuyée sur les services spécifiques d'un hébergeur, plus la sortie est longue et coûteuse. Le droit a récemment allégé cette charge, sans la supprimer : on verra que le règlement européen sur les données interdit désormais au prestataire d'entraver le départ de son client et, bientôt, de le lui facturer. Il ne fait pas pour autant le travail de la migration.

Un cas le montre, à l'échelle d'un projet national. La plateforme nationale des données de santé, créée en 2019 pour rassembler les données du système de santé au service de la recherche, a été hébergée chez Microsoft dès sa mise en service. En février 2026, le gouvernement a annoncé sa sortie de cet hébergeur et lancé une mise en concurrence. Il a officialisé en avril le choix d'un opérateur français, certifié pour l'hébergement des données de santé, dont la qualification par l'Agence nationale de la sécurité des systèmes d'information (ANSSI) était encore en cours. La bascule est attendue entre la fin de 2026 et le début de 2027. La dépendance aura donc duré près de sept ans, malgré les débats publics et les recours contentieux qui l'ont accompagnée dès l'origine. Pour autant, le constat ne juge pas le choix initial, qui s'expliquait par l'état de l'offre et par l'urgence; la commission d'enquête de l'Assemblée nationale relève d'ailleurs que l'offre française a sensiblement progressé depuis. Il établit seulement qu'une sortie, même voulue et engagée, prend des années et qu'une dépendance dont on n'a pas chiffré la sortie au moment d'y entrer dure toujours plus longtemps que prévu. L'interconnexion, lorsque personne n'en chiffre la sortie, glisse ainsi vers la dépendance.

2.6 Les détroits de la diffusion

Le canal est la voie par laquelle le fichier est mis à disposition de son lecteur. On y distingue deux niveaux, que la métaphore réunit sous l'image des routes maritimes. Le premier est physique : les réseaux de l'internet, et d'abord les câbles sous-marins par lesquels passe la quasi-totalité des échanges de données entre continents. Ce sont les routes et les câbles de la métaphore. Le second est celui des plateformes par lesquelles le fichier est publié, et qu'il faut emprunter pour l'obtenir. Ce sont les détroits, passages obligés que tient celui qui les contrôle. Ce peut être une plateforme de publication, comme data.gouv.fr, le site d'une institution ou le serveur d'un prestataire. Ce peut être une forge, plateforme où l'on dépose et partage le code des logiciels, comme GitHub ou

GitLab. Ce peut être enfin une interface de programmation (en anglais *API*), par laquelle un logiciel interroge directement une base, comme celles qui permettent de consulter le répertoire des entreprises ou la base des adresses. Qui tient le détroit tient le passage. La marchandise peut appartenir à l'armateur ; la route appartient à celui qui l'entretient, et celui-ci impose ses conditions à quiconque l'emprunte.

Le premier niveau n'est pas une image : la France a fait de la maîtrise des fonds marins, où courent ces câbles, un enjeu de sa stratégie de défense. Mais les détroits qui comptent pour notre sujet sont ceux du second niveau, et ils sont juridiques plus que physiques.

Le premier est celui des plateformes qui diffusent les modèles et les ensembles de données d'intelligence artificielle. La principale d'entre elles, Hugging Face, est devenue la bibliothèque commune de l'intelligence artificielle. Chercheurs, entreprises et administrations y déposent et y téléchargent des modèles et des ensembles de données, comme les développeurs le font pour le code sur GitHub. Qui veut publier un modèle ou un corpus pour qu'il soit trouvé et réutilisé passe presque nécessairement par elle. Fondée par des entrepreneurs français, elle est aujourd'hui une société américaine, dont la maison mère a pris la présidence de la filiale française. Le 3 septembre 2026, NVIDIA a annoncé un accord pour l'acquérir.

Le deuxième est celui des forges de code. La plus utilisée au monde, GitHub, appartient à Microsoft. La Cour des comptes y publie son outil d'analyse financière, l'équipe interministérielle chargée du logiciel libre y publie ses projets, et Euro-Office lui-même y a déposé son code. Le recensement public des dépôts de code des administrations en compte des milliers, sans que l'on sache sur quelles forges ils se répartissent. Pourtant, des solutions indépendantes existent : forges associatives européennes, comme Framagit, tenue par l'association française Framasoft, ou forges hébergées par les administrations elles-mêmes, comme celle que le ministère de l'éducation nationale a développée. La même équipe interministérielle utilise d'ailleurs aussi une forge indépendante. Le choix du détroit est donc possible ; il n'est simplement pas exercé de manière systématique.

Le troisième cas renverse la situation et montre qu'un canal peut fermer ce que la licence ouvre. Un référentiel des nomenclatures de ressources humaines de l'État est publié sous licence ouverte, ce qui autorise quiconque à le réutiliser. Pourtant, l'interface qui permet de l'interroger n'est accessible que depuis le réseau interne de l'administration. Ainsi la donnée est ouverte en droit, et le passage fermé en fait. Le connaissance autorise la livraison, mais le détroit n'est ouvert qu'aux navires de la flotte de l'État.

2.7 La persistance des défauts

Le poste de consommation est le dernier étage de notre raisonnement, celui où la cargaison doit être déchargée et lue : le logiciel de l'utilisateur, son ordinateur ou son téléphone, et désormais l'assistant d'intelligence artificielle qui lit pour lui. Dans notre métaphore, il est le quai de destination. Un quai ne décharge que ce que le navire apporte : les défauts nés en amont ou apparus pendant le voyage reparaissent fidèlement en aval.

En septembre 2026, l'outil expérimental permettant aux assistants d'intelligence artificielle de consulter la plateforme nationale des données publiques a été interrogé sur le répertoire des élus. Il a livré quatre informations erronées. Trois d'entre elles venaient des descriptions saisies par le producteur de la donnée, que l'outil avait fidèlement relayées ; une seule lui était propre. Ainsi, le quai d'arrivée n'avait pas détérioré la cargaison : elle y a été livrée telle qu'elle avait été expédiée.

C'est aussi à cet étage que se trouve ce qu'on pourrait appeler la maison du régulateur : les outils avec lesquels les autorités qui édictent les normes travaillent elles-mêmes. On y reviendra.

2.8 La licence égarée

La licence n'est pas un huitième étage. Elle porte sur le contenu, puisqu'elle dit ce que l'on peut faire des données ; mais elle voyage avec le contenant et accompagne la donnée à chacun des sept étages. Si le fichier ne la porte pas, elle se perd en route. Dans notre comparaison, elle est le connaissance maritime déjà évoqué. Une cargaison sans connaissance n'est pas perdue, mais celui qui la reçoit ne sait ni à quel titre il la détient, ni ce qu'il a le droit d'en faire.

En septembre 2026, plus d'un quart des jeux de données publics de data.gouv.fr étaient mis à disposition sans licence valide. Soit la licence avait été déclarée « non spécifiée », soit la case avait été laissée vide. Des jeux publiés par des organisations certifiées, dont les services du Premier ministre et les ministères économiques et financiers, en font partie. Dès lors, le réutilisateur ne sait pas, à la lecture de la ressource, dans quelles conditions il la reçoit. Le droit commun lui permet, en principe, de réutiliser librement une information publique, une licence n'étant obligatoire que si la réutilisation est soumise au paiement d'une redevance. Mais rien, sur la ressource, ne le lui dit : il doit reconstituer ce régime par lui-même, ce que la plupart ne feront pas.

Le connaissance le plus exigeant réserve à son tour une surprise. Une licence employée pour certaines données publiques, l'ODbL (*Open Database License*), impose ce que l'on appelle le partage à l'identique. Quiconque diffuse une base de données dérivée de la base d'origine doit la mettre à son tour à disposition dans les mêmes conditions. L'une de

ses clauses, la moins connue, va plus loin. Celui qui utilise publiquement une base dérivée, ou un produit fabriqué à partir d'elle, a une obligation envers ceux qui le reçoivent. Il doit leur offrir soit la base dérivée entière, soit la description des modifications apportées à la base d'origine, soit le moyen de les reproduire.

Que devient cette obligation lorsqu'une intelligence artificielle est entraînée sur des données placées sous cette licence? Le modèle obtenu est-il une base dérivée? Aucun texte ne le dit. La lecture la plus détaillée dont on dispose provient des lignes directrices d'une fondation privée britannique, celle qui gère la grande carte collaborative du monde, OpenStreetMap. Elle les a écrites pour ses propres données, et elles qualifient de base dérivée un ensemble d'entraînement qui reprend une part substantielle de la base. Or cette lecture est discutée au sein même de sa communauté. Elle ne lie pas les autres titulaires de données placées sous la même licence. Aucune autre interprétation publiée n'ayant toutefois été trouvée, elle est la seule lecture disponible. Une licence employée pour des données publiques françaises est ainsi lue, sur son point le plus neuf, à travers l'avis d'une fondation privée étrangère. Le connaissance rejoint ici la carte : comme le classement par le sens, la portée de la licence dépend d'une lecture élaborée ailleurs, par un acteur que personne en France n'a choisi et qui n'en répond devant aucune autorité française.

Au terme du voyage, la chaîne est complète. Chacun de ses étages a son gardien, sa dépendance et son coût de sortie, et chacun peut être choisi ou subi. Reste à savoir lesquels le droit a entrepris de régler.

3 Une loi de surface

La chaîne vient d'être décrite telle qu'elle est. Elle doit maintenant être relue telle que le droit l'appréhende. La méthode est simple : prendre les sept étages un à un et demander pour chacun quel texte le régit, avec quel degré de contrainte et avec quels moyens d'en contrôler l'application. Le résultat est net. Un seul étage, l'hébergement, est entouré d'une digue de textes complète, de la loi à l'arrêté, de la certification à la procédure de dérogation. Un deuxième, le format des fichiers, reçoit des encouragements et une règle précise que ses propres auteurs n'appliquent pas. Les autres ne sont régis que marginalement, ou pas du tout. La digue existe et elle est solide ; elle ne protège qu'un étage sur sept.

Dans notre métaphore, c'est un port où l'on inspecte avec minutie le pavillon de chaque navire, son registre d'immatriculation et le droit dont relève son armateur, tandis que nul inspecteur ne regarde la coque, la carte ni la route. Au-delà du pavillon, la chaîne ressemble à la haute mer. La convention des Nations unies sur le droit de la mer, signée à Montego Bay en 1982, y proclame la liberté de navigation et interdit à tout État de s'en approprier une partie. Elle soumet chaque navire à la seule loi de l'État dont il bat le pavillon ; les autres États n'y peuvent intervenir que dans des cas limitativement énumérés, comme la piraterie ou la traite des esclaves. Les étages du format, du modèle et du canal forment une haute mer de cette sorte. Aucune autorité publique n'y exerce de souveraineté, et seule s'y applique la loi que s'est donnée l'éditeur du logiciel, du modèle ou de la plateforme : ses propres réglages et ses propres conditions d'utilisation. Rien n'y est illégal ; tout y est laissé au choix de ceux qui construisent les navires.

3.1 Le contrôle du pavillon

L'hébergement des données publiques a fait l'objet, en moins de huit ans, d'une construction juridique continue et de plus en plus exigeante. Une première circulaire du Premier ministre, en 2018, a posé la doctrine d'utilisation de l'informatique en nuage par l'État. L'informatique en nuage, que l'on appelle couramment le *cloud*, consiste à recourir à des serveurs distants, loués à un prestataire, plutôt qu'à des machines installées chez soi. Une deuxième, en 2021, dite « cloud au centre », en a fait le mode d'hébergement de référence des administrations. Elle a posé la règle devenue emblématique, que l'on désigne par son numéro, R9. Lorsqu'un système manipule des données d'une sensibilité particulière, l'offre commerciale retenue doit impérativement disposer de la qualification délivrée par l'ANSSI, dite SecNumCloud, ou d'une qualification européenne d'un niveau au moins équivalent. Elle doit en outre être immunisée contre toute réglementation extracommunautaire. Une troisième circulaire, en 2023, a délimité plus précisément ce que recouvrent les données d'une sensibilité particulière et encadré les demandes de dérogation.

Puis le législateur a donné à cette exigence force de loi. La loi de 2024 visant à sécuriser et à réguler l'espace numérique en a fait une obligation pour l'administration de l'État, ses opérateurs et certains groupements d'intérêt public. Ceux-ci doivent veiller à ce que leur prestataire d'hébergement protège les données sensibles contre tout accès d'autorités étrangères que le droit européen ou national n'autoriserait pas. Un décret de 2026 en a fixé l'application. Il dresse la liste des six groupements concernés, parmi lesquels l'Agence du numérique en santé et le groupement chargé de la modernisation des déclarations sociales. Il renvoie à un référentiel technique élaboré par l'ANSSI, dont le respect doit être attesté par une qualification nationale ou par une certification européenne d'un niveau au moins équivalent. Un arrêté du mois d'août 2026 a approuvé la version en vigueur de ce référentiel, qui est ainsi devenu opposable. Enfin, le décret organise une procédure de dérogation encadrée : l'administration qui ne peut s'y conformer doit la solliciter auprès de son ministre, avec l'aval du Premier ministre, et pour une durée limitée. Loi, décret, arrêté, référentiel technique, qualification par une autorité nationale, procédure de dérogation motivée : c'est la pyramide normative entière.

Une enceinte complète n'est pas pour autant infranchissable. L'été 2026 l'a rappelé. La direction générale des finances publiques (DGFIP) a reconnu en août 2026 qu'un accès illégitime avait permis de consulter et d'extraire les données de 678 000 particuliers et professionnels. L'attaque n'a pas percé l'hébergement : elle est entrée avec les identifiants usurpés d'un agent et d'un tiers habilité. Or le pavillon était en règle. La brèche était ailleurs : dans l'accès des personnes au système, un point que la chaîne des sept étages ne décrit pas et que la pyramide des textes ne vise pas davantage. L'exemple ne dit pas que la protection de l'hébergement est inutile ; il dit qu'elle ne couvre que l'étage qu'elle voit.

Cet étage est aussi le seul qui ait donné lieu à un contentieux soutenu. L'hébergement de la plateforme nationale des données de santé a occupé la Commission nationale de l'informatique et des libertés (CNIL), puis le Conseil d'État. Celui-ci a jugé en mars 2026 que l'autorisation donnée par la CNIL portait sur des données hébergées en France et n'emportait aucun transfert de données de santé vers les États-Unis. Les données techniques susceptibles d'être vues depuis l'étranger étaient, selon lui, couvertes par des garanties contractuelles. Ce contentieux prolongeait un arrêt retentissant de la Cour de justice de l'Union européenne, dit *Schrems II*, qui a invalidé en 2020 le cadre sur lequel reposaient alors les transferts de données personnelles vers les États-Unis.

L'hébergement fait même l'objet de négociations internationales. Devant la commission d'enquête de l'Assemblée nationale, le délégué général du Cigref, qui fédère les grandes entreprises utilisatrices du numérique, a raconté que le projet européen de certification des services d'informatique en nuage comportait, jusqu'en 2023, un niveau de sécurité renforcé. Ce niveau, qui intégrait des critères de protection contre les lois étrangères, a

disparu du projet l'année suivante. Il l'a attribué aux pressions exercées par les États-Unis.

Ce constat ne signifie pas que le droit en fasse trop pour l'hébergement. Le risque qu'il traite est réel, et l'exemple de la plateforme nationale des données de santé a montré ce qu'il en coûte de le négliger. Le constat donne en revanche la mesure de ce qu'un étage reçoit lorsque le droit le voit. Il reçoit des textes à chaque niveau de la hiérarchie des normes, un organe technique pour les appliquer, une autorité pour certifier, un juge pour trancher, et jusqu'à des négociations internationales. Les autres étages de notre comparaison doivent être mesurés à l'aune de celui-ci.

3.2 Le format, entre vœu et lettre morte

L'étage du format n'est pas ignoré du droit français. Il est nommé dans la loi, ce qui rend son traitement révélateur.

La loi pour une République numérique de 2016 dispose que les administrations « veillent à préserver la maîtrise, la pérennité et l'indépendance de leurs systèmes d'information » et qu'elles « encouragent l'utilisation des logiciels libres et des formats ouverts ». Le texte contient le mot de pérennité, l'horizon même donné plus haut à la doctrine du contenant. Il le place sous le verbe « veiller ». Il place en revanche les formats ouverts sous le verbe « encourager ».

La loi de 2024 emploie le même verbe : les administrations « veillent » à ce que leur hébergeur protège les données. Mais elle est suivie d'un décret, d'un référentiel, d'une qualification et d'une procédure de dérogation. Le texte de 2016, lui, n'est suivi de rien. Aucun décret n'en précise la portée, aucun organe n'en mesure l'application, aucune procédure n'en sanctionne la méconnaissance. Ainsi le même verbe produit-il une obligation outillée pour l'hébergement, le pavillon de notre métaphore, et seulement un vœu pour le format des fichiers, sa coque.

Un référentiel complète le dispositif : le référentiel général d'interopérabilité, qui fixe les normes et les formats à employer pour chaque usage. Sa portée juridique n'a rien de symbolique, puisque l'ordonnance de 2005 qui le fonde prévoit qu'il détermine les normes et les standards que les autorités administratives doivent utiliser. Pourtant, il est toujours présenté comme en vigueur dans une version qui remonte à plus de dix ans, et son application n'est ni mesurée ni sanctionnée. Le rapport de la commission d'enquête de l'Assemblée nationale sur les dépendances numériques, publié en juillet 2026, en montre la portée pratique. La rapporteure y suggère, parmi les propositions qui lui sont propres et se distinguent des recommandations adoptées par la commission, de rendre obligatoire l'application de ce référentiel dans les marchés publics. Elle observe que les administrations disposent déjà des référentiels nécessaires et que la tâche est

désormais d'en imposer le respect. Proposer de rendre obligatoire, c'est constater que la règle ne l'est pas en pratique.

On pourrait penser que le défaut tient à la mollesse du droit et qu'un texte précis et impératif suffirait. L'expérience dément cette hypothèse, car ce texte existe. L'arrêté de 2017 évoqué en introduction, celui dont les services du Premier ministre ne respectent pas la première ligne, en fournit la preuve. Pour les données des subventions publiques, il fixe un modèle de fichier qui impose, dès la première ligne de son annexe, l'encodage universel et recommande la virgule pour séparer les colonnes. La règle est obligatoire et précise; elle porte sur le tirant d'eau même de la péniche. On a vu ce qu'il en est de son application par les services de son propre signataire.

L'histoire de ce texte enseigne autre chose encore. Or, en France, la virgule sépare aussi les décimales. Un même fichier ne peut guère employer la virgule à la fois pour séparer les colonnes et pour séparer les euros des centimes sans que les logiciels s'y perdent. L'équipe de l'État chargée d'appliquer le texte a donc retenu en pratique le point pour les décimales. Elle a indiqué expressément que ce choix était fait « contrairement à la lettre de l'arrêté » et que l'arrêté allait être modifié. Pour que le modèle officiel fonctionne, ceux qui étaient chargés de l'appliquer ont donc dû s'écarter de sa lettre. Pourtant, la modification n'est jamais intervenue : la version en vigueur est toujours celle d'origine. Depuis près de neuf ans, la règle écrite et la règle appliquée divergent, sur un point technique élémentaire, au sein même de l'administration qui a produit l'une et l'autre. Chaque décision se défend; le choix pratique était même sensé. C'est leur accumulation qui fait qu'un réutilisateur scrupuleux, suivant à la lettre le Journal officiel, produirait un fichier que le modèle officiel jugerait non conforme.

La coque est ainsi régie de deux manières, qui aboutissent au même résultat. Là où la règle est générale, elle encourage. Là où elle est précise, elle n'est ni appliquée ni tenue à jour par l'administration même qui l'a édictée. Dans les deux cas, aucune contrainte n'agit au moment où le fichier est produit, moment précis où le réglage par défaut décide de son encodage. La norme perd sa force à mesure qu'elle approche du paramètre qu'elle prétend gouverner.

Ce qui passait pour un détail technique en 2017 ne l'est plus. Les données publiques sont désormais ouvertes par principe, et elles sont lues, agrégées et réutilisées en masse, de plus en plus souvent par des programmes plutôt que par des personnes. Un défaut d'encodage qu'un lecteur corrigeait d'un coup d'œil se propage aujourd'hui sans que personne le voie.

3.3 Le choix allemand

L'Allemagne est allée plus loin qu'aucun autre grand État de l'Union dans le choix explicite d'un format, et y est venue par étapes. Au printemps 2025, le conseil de planification informatique (*IT-Planungsrat*), organe commun de l'État fédéral et des Länder, a pris un engagement. Les formats ouverts, et en particulier le format OpenDocument, devaient devenir d'ici 2027 le standard de l'échange de documents dans l'administration. À l'automne suivant, il a précisé que les administrations des Länder prendraient en charge ce format progressivement à partir du 1^{er} janvier 2027, l'expérience devant être évaluée en 2028. Au printemps 2026, enfin, il a adopté un socle technique commun à toute l'administration allemande. Il en a arrêté la liste des standards, couche par couche, de l'infrastructure jusqu'à l'intelligence artificielle, et les a rendus obligatoires pour les solutions de ce socle. Parmi eux figurent le format OpenDocument et le PDF accessible aux personnes handicapées, aux côtés de formats d'échange de données comme le CSV.

Saluée comme un tournant, la décision en est bien un. Mais une partie de la presse a fusionné les trois décisions pour n'en restituer qu'une seule, et en a tiré l'obligation, pour tous les documents officiels, de n'utiliser que ces deux formats, avec une échéance couperet. Or les textes sont plus mesurés. Dans la décision de 2026, les standards sont obligatoires pour les solutions du socle commun, et non pour l'ensemble des documents administratifs. L'État fédéral, les Länder et les communes s'y engagent seulement à « s'efforcer » d'utiliser ces solutions dans leurs nouveaux développements. Les annexes ne fixent aucune échéance, et la décision n'organise ni contrôle ni sanction. Les engagements de 2025, eux, portent bien une date, mais sous la forme d'une prise en charge progressive par les Länder, assortie d'une évaluation ultérieure. L'écart n'est imputable à personne en particulier. Il illustre pourtant, au sujet d'une norme sur les formats, le mécanisme décrit plus haut au sujet des fichiers : l'étiquette voyage plus vite que la cargaison qu'elle décrit, et aucun contrôle ne la vérifie à l'embarquement.

Le même texte contient un passage que la presse a laissé de côté et qui importe davantage à notre sujet. Pour chaque couche, la décision énumère aussi les normes qui restent à définir. Dans la couche de l'intelligence artificielle, elle cite un format pour intégrer des sources documentaires aux systèmes qui consultent des textes avant de répondre. C'est ce que l'on appelle la génération augmentée par recherche (en anglais *retrieval-augmented generation*, ou RAG), soit l'usage même des bibliothèques classées par le sens décrit plus haut. Elle mentionne aussi un format d'échange des modèles de langage eux-mêmes. Dans une autre couche, elle prévoit un format pour stocker les coordonnées que produisent ces modèles. Ainsi l'administration allemande a-t-elle vu l'étage de la carte. Elle ne l'a pas réglé, elle l'a inscrit comme une lacune à combler. C'est le seul texte officiel rencontré au cours de cette enquête qui identifie ainsi cet étage.

Pour autant, le mandat fédéral n'empêche pas les trajectoires divergentes. Au moment où la décision était adoptée, la Bavière négociait avec Microsoft un contrat-cadre pour l'ensemble de ses administrations, évalué par la presse à près de 1 milliard d'euros sur cinq ans. Le projet a divisé publiquement son propre gouvernement. Les négociations ont été abandonnées en juin 2026. À l'inverse, le Schleswig-Holstein a conduit la migration la plus avancée d'Allemagne. Selon son gouvernement, près de quatre postes de travail sur cinq de l'administration, hors administration fiscale, fonctionnaient à la fin de 2025 sous une suite bureautique libre. L'économie de licences dépasserait nettement l'investissement consenti. Le même gouvernement indique que le cinquième restant demeure lié à Microsoft, parce que des logiciels métier en dépendent techniquement. La dépendance résiduelle tient donc moins au choix du poste de travail qu'aux outils d'autres chantiers navals, qui attendent des coques d'un certain gabarit.

L'Estonie, souvent citée comme l'un des États les plus numérisés d'Europe, a suivi une voie comparable par son cadre d'interopérabilité. Celui-ci fixait une liste minimale de standards ouverts, parmi lesquels le format OpenDocument, le PDF/A et le CSV, que les administrations devaient employer dans leurs échanges entre elles et avec le public. Pour leur usage interne, d'autres formats restaient permis, quoique déconseillés. Refondu en 2023, ce cadre se présente désormais comme un document évolutif plutôt que comme un texte normatif. La règle existe donc, mais elle s'arrête à la porte des services.

La Norvège offre un point de comparaison plus ancien. Dès 2009, son catalogue des standards publics rendait le format OpenDocument obligatoire pour les documents destinés à être modifiés par leur destinataire. Son règlement actuel impose aux administrations un format d'archivage pour les documents envoyés aux citoyens et aux entreprises ; il les oblige aussi à savoir recevoir les formats bureautiques courants, quels qu'ils soient. Il ne permet d'y déroger qu'en cas de charge particulièrement disproportionnée et à la condition d'en informer sans délai l'agence nationale du numérique, en motivant la dérogation. C'est, dans toute la comparaison, le seul dispositif qui assortisse une règle de format d'une procédure : la dérogation est possible, mais l'administration doit la déclarer.

Dans notre métaphore, l'Allemagne a fixé les plans de la flotte commune et désigné les cartes qui restent à dresser, sans placer d'inspecteur au chantier. La Norvège a ouvert un registre des exceptions.

3.4 L'interopérabilité en question

L'Union européenne a choisi une voie différente. Un règlement de 2024, dit « pour une Europe interopérable », vise les institutions européennes et les administrations nationales qui s'appêtent à créer, ou à modifier sensiblement, une exigence obligatoire touchant des services publics numériques transfrontaliers. Avant de l'adopter, elles doivent évaluer son

effet sur la capacité des systèmes à échanger entre eux. Le règlement ne prescrit aucun format. Il prescrit une question : quel sera l'effet de cette décision sur la capacité des systèmes à communiquer entre eux ?

Le droit de la commande publique procède de la même manière. La directive européenne sur les marchés publics interdit, sauf justification tenant à l'objet du marché, d'exiger une marque, une origine ou un procédé particulier et impose d'accepter les produits équivalents. Elle protège la concurrence entre fournisseurs. Elle ne dit rien, en revanche, du format que produira le logiciel acheté.

On aurait tort de voir dans cette retenue une faiblesse. En imposant l'évaluation plutôt que le résultat, le règlement fait à son échelle ce que la conclusion recommandera au dirigeant. Il pose la question du contenant au moment de la décision, plutôt que de désigner une fois pour toutes le bon format. Toutefois, il le fait dans un champ étroit, celui des services publics transfrontaliers et des seules exigences obligatoires. Le fichier qu'une administration nationale publie pour ses propres administrés n'y entre pas.

Dès lors, pour le document que l'on produit, seul l'État impose un format dans l'ordre juridique européen. L'Union oblige à se poser la question ; elle ne prescrit un format qu'au moment où l'on quitte un prestataire, comme on va le voir. L'Allemagne et la Norvège y répondent pour leurs administrations. La France y a répondu une fois, par l'arrêté sur les subventions, dont on a vu le sort. Pour tout le reste, la réponse appartient à l'éditeur du logiciel, qui en fixe le réglage par défaut.

3.5 Le droit de reprendre sa cargaison

L'Union a pourtant franchi, sur un point, la limite qu'elle s'impose ailleurs. Le règlement européen sur les données, dit *Data Act*, adopté en 2023 et applicable depuis le 12 septembre 2025, consacre un chapitre entier au changement de fournisseur de services de traitement de données. La catégorie est large. Elle couvre la location de serveurs à distance, les plateformes de développement et les logiciels utilisés en ligne sur abonnement, que l'on désigne par le sigle anglais SaaS (*software as a service*). La messagerie d'une entreprise, son logiciel de comptabilité ou le tableur partagé de ses équipes en relèvent le plus souvent. Le règlement s'applique directement, sans transposition, aux entreprises comme aux administrations.

Il interdit au fournisseur d'entraver le départ de son client, que l'obstacle soit contractuel, technique ou commercial. Il l'oblige à décrire à l'avance les procédures, les formats et les limites de la sortie. Il organise la disparition des frais : réduits aux seuls coûts du changement depuis janvier 2024, ils seront interdits à compter du 12 janvier 2027. La France a pris les devants pour une partie d'entre eux, puisque les frais de transfert de

données facturés à l'occasion d'un changement de fournisseur y sont plafonnés à zéro euro depuis la fin de 2025. Surtout, tant qu'aucune norme commune n'a été publiée pour le type de service concerné, le fournisseur doit exporter, à la demande du client, toutes les données exportables, dans un format structuré, couramment utilisé et lisible par machine. Le texte est lui-même en mouvement : une proposition de la Commission, encore en discussion, dispenserait de ces obligations certains prestataires, pour les contrats conclus avant le 12 septembre 2025. Pour les données personnelles, le RGPD reconnaissait déjà depuis 2018 aux personnes concernées un droit de même facture, celui de la portabilité.

Voilà donc un texte européen qui impose un résultat et qui nomme le format. Le constat nuance la « loi du visible » ; il ne la dément pas, et la manière dont le texte saisit le format en donne la raison.

Le règlement saisit d'abord le format à un moment précis : celui du départ, et à la demande du client. Il ne dit rien du fichier au moment où il est produit, là où le réglage par défaut décide de son encodage. Dans notre métaphore, le droit garantit à l'armateur qu'il pourra reprendre sa cargaison le jour où il change de compagnie. Il ne regarde pas la coque dans laquelle elle a voyagé jusque-là.

Il le saisit ensuite par une formule dont chaque terme se mesure à l'aune des sept questions posées plus haut. Un fichier CSV répond sans peine aux trois exigences, et rien dans la formule n'interdit qu'il soit écrit dans l'ancienne table de caractères de Windows, comme celui des subventions qui ouvre cet article. La formule ne dit rien de l'encodage, rien de la gouvernance du format, rien de sa pérennité. Elle garantit que la cargaison sera rendue ; elle ne garantit pas qu'elle sera encore lisible dans vingt ans.

Enfin, il laisse hors du portage ce que le fournisseur a lui-même apporté. Rien ne l'oblige à transmettre les éléments protégés par la propriété intellectuelle ou le secret des affaires. L'obligation la plus exigeante, celle de faciliter chez le nouveau fournisseur un fonctionnement équivalent, ne pèse que sur la location d'infrastructure, et non sur les logiciels en ligne. On reprend ses données ; on ne reprend ni l'outil qui les rendait utiles, ni la manière dont il les classait. Qu'un fonds documentaire ait été classé par un modèle appartenant au fournisseur, et l'on pourra en exporter les textes et leurs coordonnées de sens. On n'exportera pas avec eux la capacité de les interroger : la clef sémantique décrite plus haut reste chez celui qui la détient.

Le mérite du texte n'en est pas moins considérable. Il fait de la sortie un droit, là où elle n'était qu'une négociation, et il vaut pour l'entreprise autant que pour l'État. Il rappelle aussi au dirigeant qu'un droit de sortie ne s'exerce utilement que s'il a été préparé : savoir ce que l'on détient, chez qui, dans quel format, et ce qu'il faudrait pour le reprendre. Le droit garantit désormais qu'on pourra quitter le port. Il ne dit toujours pas dans quelle

coque la cargaison y était entrée.

3.6 Un code public sans port d'attache

Les administrations développent elles-mêmes de nombreux logiciels : un outil de calcul des impôts, une application de gestion, un site de démarches. Le code de ces logiciels (le texte des instructions qui les font fonctionner) est considéré depuis la loi pour une République numérique comme un document administratif. Il est donc communicable à qui le demande et, dans certaines conditions, doit être publié. Une circulaire du Premier ministre en a fait un axe de la politique publique de la donnée. Le droit a vu la cargaison et en a organisé l'ouverture.

Mais il n'a pas vu la route. Aucun texte ne dit où un code public doit être publié, sur quelle forge, sous quel droit, avec quelles garanties de pérennité. On a vu ce qu'il en résulte : la Cour des comptes publie son outil d'analyse financière sur une forge appartenant à une entreprise américaine. Le dépôt de référence du modèle de fichier des subventions, celui-là même qui fixe la pratique du séparateur décimal, y est déposé lui aussi. La commission d'enquête de l'Assemblée nationale a vu ce détroit. Sa recommandation n° 17 tend à soutenir et à valoriser la forge publique développée par le ministère de l'éducation nationale. Mais c'est une recommandation de soutien, non une règle de lieu. L'obligation d'ouvrir est posée ; l'obligation de choisir le passage ne l'est pas.

3.7 La licence européenne en suspens

Hors la sortie d'un prestataire, le connaissance est le seul point de la chaîne où le droit européen impose un résultat. Le droit européen connaît une catégorie de données publiques dites « de forte valeur ». La directive de 2019 sur les données ouvertes les définit comme celles dont la réutilisation est associée à d'importantes retombées positives pour la société, l'environnement et l'économie. Pour elles, un règlement européen oblige à une licence qui laisse entièrement libre la réutilisation. Il admet pour seule condition éventuelle la citation de la source, à moins que la licence ne soit encore moins restrictive. Parmi ces données figurent celles de la qualité de l'air.

Or la surveillance de la qualité de l'air est confiée en France à des associations agréées, dont plusieurs publient leurs données sous la licence ODbL évoquée plus haut, ou sans licence déclarée. Il est donc difficile de soutenir qu'une licence de partage à l'identique soit équivalente, ou moins restrictive, qu'une licence qui n'exige que la citation de la source. L'obligation de repartager à l'identique est une restriction que cette dernière ne comporte pas. Faut-il en conclure que l'obligation européenne pèse sur ces associations elles-mêmes, ou seulement sur l'État, qui détient les données qu'elles lui transmettent ? Aucune position officielle, ni de la Commission européenne, ni de l'administration fran-

çaise, ne l'a tranché. La réutilisation de ces données, qui est la raison d'être de leur ouverture, en dépend pourtant. Le connaissance que portent ces données n'est donc pas celui que le droit européen exige pour la même cargaison, sans que personne ait décidé qu'il en irait ainsi.

3.8 Modèles et plateformes : le silence des textes

Il reste les étages sur lesquels aucune norme ne porte. Aucun texte français ne régit le modèle qui classe par le sens un ensemble de documents publics, ni la dépendance qu'il installe, ni le coût d'en sortir. Le règlement européen sur les données n'y supplée pas : il organise la sortie d'un prestataire, non la dépendance à un modèle que l'organisation exploite elle-même. Aucun ne régit le choix du canal par lequel un tel ensemble est diffusé. Le règlement européen sur l'intelligence artificielle encadre les systèmes selon le niveau de risque qu'ils présentent et impose des obligations aux fournisseurs des grands modèles. En revanche, il n'a pas pour objet la dépendance d'un fonds documentaire public à l'égard du modèle qui l'indexe. Par sa recommandation n° 18, la commission d'enquête de l'Assemblée nationale demande que les modèles présentés comme ouverts le soient réellement, au besoin par des actions en justice. C'est une question sur le modèle lui-même, non sur ce que le modèle sert à retrouver.

Le seul texte officiel qui voie cet étage est, on l'a dit, la décision allemande, qui le range parmi les normes restant à définir. La carte est donc vue, par un seul État, comme une lacune ; aucun ne l'a encore comblée.

3.9 Une mesure sans conséquence

Entre le chantier et le quai, il existe pourtant un lieu où passent un grand nombre de fichiers publics et où leur contenant pourrait être constaté : la plateforme nationale des données publiques, data.gouv.fr. Une mesure conduite pour cet article en septembre 2026, sur près de 400 000 ressources, permet d'en décrire le fonctionnement.

Le premier problème tient à la déclaration. Le format d'une ressource y est déclaré par son producteur lui-même, en texte libre, sans liste de choix. On relève ainsi plus de 2 000 manières différentes de l'écrire. Le producteur déclare donc sa marchandise sans nomenclature.

La plateforme ne se contente pas de cette déclaration. Un programme de contrôle automatique, ce que les informaticiens appellent un robot, télécharge les fichiers et en vérifie le type réel. Là où la comparaison est possible, déclaration et réalité concordent dans plus de 96 % des cas. La fausse déclaration de type est donc marginale. Les fichiers déposés directement sur la plateforme sont les plus fidèles, avec des écarts presque nuls ; les écarts

se concentrent sur les fichiers que la plateforme récupère automatiquement sur d'autres sites, ce qui rejoint l'étage du canal. Le deuxième problème tient aux limites de ce contrôle. La comparaison n'est possible que pour un peu plus d'un tiers des ressources. Un quart d'entre elles n'a jamais été analysé. Plus de 7 % sont inaccessibles. Et sur un échantillon de fichiers tableurs dont l'encodage a été mesuré, environ un sur seize n'était pas dans l'encodage attendu. Ce chiffre vaut pour cet échantillon et non pour l'ensemble.

Le troisième problème, le plus sérieux, tient à ce que la plateforme affiche. D'après son code publié, ce même robot détecte l'encodage des fichiers pour pouvoir les lire lui-même, mais il ne l'affiche pas. Il conserve la déclaration du producteur ; le site ne montre au lecteur que cette déclaration. Le choix est délibéré et se défend : l'hébergeur ne réécrit pas ce que déclarent les producteurs, qui en répondent. Mais il a une conséquence. Dans notre métaphore, le douanier inspecte la coque, mesure son tirant d'eau, consigne le résultat dans son registre, puis laisse afficher à quai la déclaration d'origine. La plateforme a les moyens d'émettre une réserve ; elle ne l'émet pas.

On ne saurait en faire grief à la plateforme, qui mesure plus et mieux que quiconque. L'argument vaut au contraire a fortiori : si le port le mieux équipé ne tire pas de conséquence visible de ce qu'il mesure, ceux qui ne mesurent rien ne le peuvent pas.

3.10 Faites ce que je dis, pas ce que je fais

Les institutions qui font la loi, qui la contrôlent ou qui mesurent la nation publient elles aussi des données. Appliquent-elles à leurs propres fichiers les règles qu'elles fixent aux autres ? Une mesure conduite par la revue le 20 septembre 2026 sur data.gouv.fr apporte une première réponse, pour cinq d'entre elles. Pour chacune, on a relevé les licences déclarées, les formats utilisés et le lieu où les fichiers sont hébergés, puis tiré au hasard jusqu'à 12 fichiers tableurs pour vérifier s'ils sont écrits dans l'encodage universel. Les échantillons sont petits : les résultats valent comme indices, non comme verdicts.

L'Institut national de la statistique et des études économiques (Insee) soigne ses fichiers. Sur près de 200 jeux publiés, presque tous portent une licence ouverte, et tous les fichiers tableurs de l'échantillon qui ont pu être lus sont dans l'encodage universel. L'Insee publie même une partie de ses données dans un format moderne, conçu pour les très gros volumes. Son seul défaut visible est administratif : pour près de 200 ressources, le format n'a pas été déclaré.

Le Conseil constitutionnel, qui publie peu, publie proprement : ses fichiers sont presque tous hébergés sur la plateforme elle-même, et tous ceux de l'échantillon qui ont pu être lus sont dans l'encodage attendu.

En revanche, le Parlement offre un tableau plus contrasté. Au Sénat, 10 des 12 fichiers

tableurs tirés au hasard ne sont pas dans l'encodage universel, mais dans l'ancienne table de caractères de Windows. Pour tout logiciel réglé autrement, le mot « Requête » en sort déformé dès la première ligne. Ses fichiers sont, pour l'essentiel, hébergés hors de la plateforme nationale. À l'Assemblée nationale, qui publie surtout des formats d'échange entre logiciels, les 2 fichiers tableurs testés sont dans le même cas que ceux du Sénat.

Enfin, les services du Premier ministre publient beaucoup : près de 180 jeux, plus de 500 ressources. Parmi ces jeux, 13 n'ont pas de licence déclarée. Parmi les fichiers, plus de 200 sont des PDF et 39 sont au format .doc, l'ancien format binaire de Word. Sur les 12 fichiers tableurs tirés au hasard, 10 ne sont pas dans l'encodage universel, ce qui rejoint le cas des subventions qui ouvre cet article.

Il n'y a là ni classement ni procès. Les institutions les plus rigoureuses sont celles dont le métier est la donnée elle-même. Les autres ont d'autres missions, et leurs fichiers portent la trace des outils avec lesquels ils ont été produits. Mais le constat a sa force. Le Parlement, qui vote les lois sur le numérique, et les services du Premier ministre, au nom duquel a été pris l'arrêté imposant l'encodage universel, sont ceux dont les fichiers le respectent le moins. Coïncidence, sans doute ; ironie, assurément.

3.11 Une digue de papier

La relecture de la couche juridique, étage par étage, dessine une figure d'une grande régularité : celle d'une loi de surface, qui ne saisit de l'iceberg que sa partie émergée. C'est la « loi du visible » posée plus haut.

Le droit est fort sur le pavillon, l'hébergement : il l'encadre par une loi, un décret, un référentiel de sécurité, une certification, une procédure de dérogation, et il l'a défendu devant les juges comme dans les négociations internationales.

Il est faible sur la coque, le format des fichiers : un vœu dans la loi de 2016, un référentiel vieux de plus de dix ans qu'une rapporteure propose de rendre obligatoire. S'y ajoute un arrêté précis sur l'encodage que ni son signataire n'applique ni son administration n'a mis à jour.

Il est procédural à l'échelle de l'Union, qui impose d'évaluer les effets d'une décision sans prescrire de format. Il ne devient prescriptif qu'en deux points : à la sortie d'un prestataire, où il exige un format lisible par machine sans dire lequel, et sur un connaissance, la licence des données de forte valeur, que des données visées ne portent pas.

Il est absent, enfin, sur la carte et sur le choix du détroit (le modèle d'intelligence artificielle qui classe les documents publics et la plateforme par laquelle ils circulent), dont il n'organise que la sortie. Ainsi, la contrainte faiblit à mesure qu'on s'éloigne de

l'hébergement.

On retrouve ainsi le paradoxe posé au début : on peut être en règle sur l'hébergement et laisser à d'autres la clef qui permet de chercher dans ses propres textes. Ce n'est pas une faute des administrations qui appliquent la norme. C'est une erreur d'objet : la norme a saisi le maillon qui lui était visible. Mais ce qui détermine réellement ce que le lecteur pourra lire et retrouver, le format, le modèle et la plateforme, échappe à son regard.

Le rapport de la commission d'enquête de l'Assemblée nationale sur les dépendances numériques offrait, par son ampleur, la meilleure occasion de vérifier cette lecture : 453 pages, nourries de 45 auditions au cours desquelles 112 personnes ont été entendues. Il confirme la « loi du visible », en la nuanciant. Ses recommandations portent pour l'essentiel sur l'hébergement et l'informatique en nuage ; elles voient aussi les forges de code et l'ouverture des modèles. Le format y entre par une audition favorable au format OpenDocument et par une proposition propre à la rapporteure. Sur plus de 400 pages, en revanche, l'encodage des fichiers, le classement des documents par les modèles et le coût de sortie de ces modèles ne sont jamais mentionnés. La durabilité n'apparaît qu'en passant, et le réglage par défaut n'est abordé que sous l'angle de la vie privée. Le regard du Parlement descend donc, lentement, du pavillon vers la coque. Mais il n'a pas encore atteint la carte. Et le seul texte officiel qui l'ait atteinte, pour la désigner comme une lacune, est étranger.

La digue de papier existe. Elle est haute, solide et bien entretenue ; elle ne protège qu'un étage sur sept.

Faut-il y voir de l'aveuglement, un défaut de compréhension, un manque de vision ou l'effet d'influences intéressées ? La revue ne prétend pas le trancher, faute d'éléments qui permettraient de le faire sans procès d'intention. Elle constate seulement un manque manifeste sur un sujet qui touche à l'essentiel, la souveraineté, et une dépendance que chacun a acceptée sans que personne l'ait choisie.

4 Le régulateur face à ses choix

Un seul cas permet de vérifier tout ce qui précède : l'État y est à la fois celui qui fixe les règles et celui qui les applique chez lui. Dans notre métaphore, c'est le porte-conteneurs : une seule traversée, et toutes les questions posées jusqu'ici réunies sur le même pont. Le cas n'est pas choisi parce qu'il serait fautif. Chaque décision qui l'a produit se défend ; plusieurs sont même excellentes. C'est leur accumulation qui produit le résultat que l'on va décrire.

4.1 Le droit et la machine

Pour qu'un assistant d'intelligence artificielle retrouve un article de loi à partir d'une question posée en langage courant, les textes doivent d'abord être préparés. Comme on l'a déjà évoqué, chaque texte est découpé en passages, puis chaque passage est traduit en coordonnées de sens : c'est la bibliothèque rangée par sujet plutôt que par ordre alphabétique. Or ce travail est long et coûteux, et chaque administration qui voulait bâtir un tel outil devait jusqu'ici le refaire pour son compte.

L'État l'a donc fait une fois pour toutes. Sous le nom de MediaTech, son équipe chargée de l'intelligence artificielle met à disposition une collection de textes publics déjà préparés pour la recherche par le sens. On y trouve le droit en vigueur (codes, lois et règlements), les décisions du Conseil constitutionnel et les fiches pratiques de Service-Public.fr. S'y ajoutent l'annuaire des administrations de l'État et celui des collectivités, le catalogue de la plateforme nationale des données publiques, des documents de la CNIL et les dossiers législatifs. Chaque passage y est accompagné de ses coordonnées de sens, calculées par un même modèle. Les fiches de présentation de plusieurs de ces ensembles les qualifient de *semantic-ready* (prêts pour la recherche par le sens). Ce terme décrit ce que l'on a appelé plus haut le « prêt à l'emploi » : un arbitrage rendu une fois, en amont, et hérité par tous ceux qui bâtiront dessus. Cet héritage est inscrit dans la construction même : on ne peut s'en défaire qu'en démolissant l'édifice pour le rebâtir.

Le service rendu est réel. Une équipe qui voulait interroger le droit en vigueur par le sens devait auparavant télécharger, nettoyer, découper et convertir elle-même des millions d'articles. Elle dispose désormais d'une collection préparée, documentée et placée sous licence ouverte, ainsi que d'un code public qui permet d'en reproduire la fabrication. Surtout, rien dans ce projet n'est improvisé ; tout y est fait à découvert. C'est ce qui rend le cas parlant : le résultat qui va être décrit ne procède d'aucune négligence.

4.2 Un modèle venu de Chine

La traduction des textes en coordonnées de sens exige un modèle : la carte marine de notre métaphore. L'État a retenu le même modèle pour toute la collection : BGE-M3. Il est publié en libre accès par l'Académie d'intelligence artificielle de Pékin (*Beijing Academy of Artificial Intelligence*, BAAI), un institut de recherche chinois à but non lucratif. Celui-ci a été créé en 2018 avec le soutien de la municipalité de Pékin et du ministère chinois de la science et de la technologie. Il se présente comme un institut de recherche et développement « d'un type nouveau », chargé de nourrir aussi bien la recherche que l'écosystème industriel de l'intelligence artificielle.

Le choix de ce modèle n'a rien d'une bizarrerie. L'État n'en a pas publié les motifs, du moins dans les documents consultés pour cet article ; ses qualités techniques, décrites par ses auteurs dans leur publication de 2024, suffisent pourtant à l'expliquer. Il traite plus de cent langues, dont le français, dans un même espace de sens. Il combine trois manières de chercher : par le sens général, par les mots exacts, ce qui compte pour retrouver un numéro d'article ou un terme juridique, et par une comparaison plus fine, mot à mot. Il accepte des textes longs, jusqu'à 8 192 unités de texte (en anglais *tokens*, des fragments de mots), soit plusieurs milliers de mots, là où beaucoup de modèles comparables s'arrêtent à 512. Il est enfin publié sous licence libre, et assez léger pour être installé sur les serveurs de l'administration elle-même. Son adoption depuis lors confirme que le choix était techniquement fondé. En septembre 2026, sa fiche sur Hugging Face faisait état de plus de 37 millions de téléchargements en un seul mois, et l'hébergeur français OVHcloud le propose dans son catalogue. Au regard de la seule technique, la France avait donc fait un bon choix, peut-être le meilleur disponible. C'est ce qui rend la question gênante : le meilleur outil disponible pour chercher dans le droit français venait d'un institut chinois, et aucun document public consulté ne dit si la question de son origine a été posée.

Deux précisions évitent tout malentendu. La première porte sur le statut de cet institut. Depuis mars 2025, il figure sur l'*Entity List*, la liste par laquelle le gouvernement américain restreint la vente de technologies américaines à certaines entités étrangères. Selon les termes de la décision américaine, l'institut aurait acquis ou tenté d'acquérir des biens d'origine américaine au soutien de la modernisation militaire chinoise. Il aurait développé à des fins de défense de grands modèles d'intelligence artificielle et des puces de calcul avancées. Cette inscription soumet à autorisation, sous une présomption de refus, les exportations vers l'institut des biens relevant de la réglementation américaine, y compris certains biens produits à l'étranger. Elle n'interdit en revanche à personne d'utiliser les modèles qu'il publie. C'est un acte de contrôle des exportations pris par un État tiers au nom de sa sécurité nationale, non un jugement sur la qualité du modèle ; cet article n'en tire aucune conclusion sur celle-ci.

La seconde précision porte sur la sécurité des données. Un modèle installé sur des serveurs que l'on maîtrise n'envoie rien à son auteur. Dès lors, il n'y a ici aucune fuite de données, et aucune n'est alléguée.

La licence sous laquelle il est publié, la licence MIT, est l'une des plus libérales qui soient. Cette gratuité peut surprendre au regard des milliards investis dans l'intelligence artificielle. Elle s'explique en partie : un modèle de ce type, qui classe des textes sans en rédiger, coûte beaucoup moins cher à entraîner que les grands modèles qui génèrent du texte. D'autres acteurs, comme Meta ou la société française Mistral AI, publient aussi certains de leurs modèles en libre accès. Mais la gratuité a aussi une fonction. Un outil gratuit, performant et largement adopté devient un standard, et le standard est une forme d'influence : il ne coûte rien à celui qui l'exerce, et il se paie plus tard, par celui qui voudrait en changer.

Ce modèle n'est d'ailleurs pas entièrement l'œuvre de son éditeur. Sa fiche technique indique qu'il a été bâti à partir de XLM-RoBERTa, un modèle publié en 2019 par l'équipe de recherche de Facebook, aujourd'hui Meta, dont l'Académie a étendu la capacité. Ainsi la carte a-t-elle été tracée sur le fond de carte d'un autre cartographe. La nationalité d'un modèle est composite : raison de plus pour tenir la nationalité pour un indicateur plutôt que pour un critère.

Qu'on imagine une administration qui rangerait toutes ses archives selon le plan de classement d'un bibliothécaire étranger. Rien ne sortirait du bâtiment, rien ne serait dérobé. Mais retrouver un dossier obligerait désormais à raisonner selon le plan de ce bibliothécaire ; le jour où l'on voudrait en changer, les archives devraient être reclassées une à une. La dépendance ne porte pas sur ce qui sort ; elle porte sur ce qui ordonne. L'espace dans lequel le droit français est cherché a été fixé par un modèle dont l'entraînement, l'évolution et les priorités sont décidés ailleurs. C'est la capture cognitive déjà décrite, sans intention de quiconque.

4.3 Un coût de sortie jamais évalué

Reclasser, ici, veut dire tout recalculer. Les coordonnées de la collection ne se lisent qu'avec le modèle qui les a produites. Pour le seul droit en vigueur, cela représente plus de 2 millions de passages et un volume de données considérable : chaque passage traîne derrière lui sa longue suite de nombres, stockée de surcroît sous une forme peu économe. Le modèle, publié en libre accès, peut certes être conservé, ce qui écarte le risque de perdre la clef. Il n'écarte pas pour autant celui de ne pouvoir en changer qu'à grand prix, que ce soit pour adopter un modèle européen ou plus performant, ou pour se plier un jour à une contrainte extérieure.

L'État n'a pas chiffré publiquement ce que coûterait ce changement, en temps de calcul, en mois de travail et en argent. Rien ne l'y obligeait : aucune règle, on l'a vu, ne porte sur cet étage. Existe-t-il, pour le français, des modèles européens d'une qualité comparable ? Des bancs d'essai publics permettent de le trancher. Pourtant, la question vient en second. La première est celle que l'on adressait plus haut au dirigeant : si l'on devait partir, que faudrait-il reprendre, en combien de temps et à quel coût ? Ce n'est pas une erreur de ne pas l'avoir posée ; c'est un choix abandonné, au sens où on l'a défini. Mais une dépendance dont on n'a jamais chiffré la sortie n'a pas été choisie. Elle a été acceptée.

4.4 Une dépendance active

L'État ne se contente pas de publier cette collection pour des réutilisateurs lointains. Il en fait usage lui-même. Albert, sa plateforme d'intelligence artificielle, donne aux administrations accès à des modèles hébergés sur une infrastructure qualifiée au plus haut niveau de sécurité français, ou sur le nuage interne de l'État. Or son catalogue propose le même modèle comme outil de classement par le sens pour les systèmes de recherche des administrations, accompagné d'un second modèle du même institut pour trier les résultats. La dépendance est donc active : l'État sert lui-même à ses propres services une carte dressée ailleurs.

La documentation de la plateforme affirme qu'aucune donnée ne sort de l'infrastructure ; rien ne permet d'en douter. C'est ce qui fait la valeur du cas. Le pavillon est irréprochable : infrastructure qualifiée, hébergement souverain, conformité à l'exigence de la loi de 2024. La carte, elle, est dressée ailleurs. Le paradoxe posé au début de cet article se vérifie ici à l'état pur.

Le rapport parlementaire sur les dépendances numériques décrit ce socle comme reposant sur des infrastructures de confiance et sur des modèles d'intelligence artificielle « open source et français ». Certes, la description est exacte pour une part du socle. Elle ne l'est en revanche qu'à moitié pour la carte, publiée sous licence libre sans être française, et la distinction n'apparaît pas dans une enquête qui a examiné le pavillon avec le plus grand soin. Ni l'État ni la commission n'ont menti ou dissimulé. La commission a porté son regard là où la norme lui avait appris à regarder. Le romancier canadien Robertson Davies l'a écrit d'une phrase : « L'œil ne voit que ce que l'esprit est prêt à comprendre. »

4.5 Entre Pékin et Washington

La collection MediaTech est diffusée depuis Hugging Face, la plateforme américaine devenue la principale place d'échange des modèles et des données d'intelligence artificielle. Mais l'État la publie aussi sur sa propre plateforme nationale des données publiques. Le passage n'est donc pas unique : si le détroit venait à se fermer, une autre route existe. C'est

une bonne pratique, à porter au crédit du projet. Elle montre qu'un choix de canal peut être réfléchi, dès lors qu'on y pense.

Pourtant, un constat demeure, sans grief pour personne. La carte vient d'un institut que les États-Unis ont placé sur une liste restrictive, et le principal canal de diffusion relève du droit américain. Deux décisions étrangères, qui ne visent pas la France et qui ne se connaissent pas l'une l'autre, encadrent ainsi la manière dont le droit français est aujourd'hui cherché par ses propres outils. La France n'est visée par aucune d'elles; elle se trouve simplement placée entre les deux. Qu'une tension commerciale conduise un jour Washington à durcir ses règles, ou que Pékin cesse de publier ses modèles en libre accès, et la collection ne cesserait pas de fonctionner du jour au lendemain. Mais son évolution, ses mises à jour et le choix de son successeur deviendraient des questions de politique étrangère.

La même scène se joue, à une autre échelle, à Bruxelles. La Commission européenne, qui rédige les règles européennes d'interopérabilité, travaille elle-même sous la suite bureautique de Microsoft. En 2024, le Contrôleur européen de la protection des données, l'autorité qui veille au respect de la vie privée dans les institutions de l'Union, lui a reproché plusieurs manquements dans l'usage qu'elle en faisait et lui a ordonné d'y remédier. Il a clos la procédure l'année suivante, après que la Commission a renégocié avec Microsoft les clauses de son contrat. Deux précisions évitent une mauvaise compréhension de l'exemple. L'affaire portait sur les données personnelles, non sur les formats. Et la Commission l'a réglée par un aménagement du contrat, non par un changement d'outil. Elle dispose pourtant, depuis plusieurs années, d'un bureau consacré aux logiciels libres et a développé son propre éditeur libre de textes législatifs. Elle connaît le sujet et en a les moyens; l'outil par défaut demeure pourtant.

À Paris comme à Bruxelles, à l'échelle d'un État comme à celle de l'Union, le schéma est donc le même. L'autorité qui édicte la norme l'applique scrupuleusement à l'étage qu'elle voit : l'hébergement français est qualifié, le contrat européen est mis en conformité. Et elle hérite, aux étages qu'elle ne voit pas, de choix faits ailleurs : une carte dressée à Pékin sur le fond de carte d'un autre, une suite bureautique réglée dans l'État de Washington, où Microsoft a son siège. Pour autant, aucune de ces autorités n'a mal agi. Chaque décision, prise isolément, se défend. C'est leur accumulation qui produit un résultat que personne n'a souhaité, selon la formule même qui fonde cette revue.

5 La pérennité du contenant

La durabilité a été posée plus haut comme l'horizon de toute la chaîne. Il reste à voir ce que le droit en dit. Il en dit davantage qu'on ne le croit ; son silence ailleurs n'en est que plus significatif. Le droit de la preuve, le droit du patrimoine et le droit européen des services de confiance, ces prestations qui garantissent l'authenticité des échanges électroniques (signature, cachet, horodatage), connaissent tous trois la question de la conservation. Mais tous trois la saisissent au même endroit : à la fin du voyage, lorsque le document entre dans un système d'archivage ou se présente devant un juge. Dans notre métaphore, c'est la douane du port d'arrivée. Seul contrôle suivi d'effet de toute la traversée, elle intervient lorsque la cargaison a déjà voyagé dans sa coque et qu'elle est arrivée à destination. Ce qui a été mal chargé au départ arrive tel quel, et la douane ne peut que le constater.

5.1 La lisibilité de la preuve

Le droit de la preuve est le premier à rencontrer le contenant, parce qu'il doit dire à quelles conditions un écrit électronique vaut un écrit sur papier. Le Code civil lui reconnaît la même force probante, à deux conditions : que l'on puisse identifier la personne dont il émane et qu'il soit établi et conservé dans des conditions qui en garantissent l'intégrité. La conservation est donc une condition de la preuve.

Le même code va plus loin lorsqu'il traite des copies : il nomme la forme. Une copie fiable a la même force probante que l'original. Elle est présumée fiable lorsqu'elle reproduit à l'identique la forme et le contenu de l'acte et que son intégrité est garantie dans le temps par un procédé dont un décret fixe les conditions techniques. Ainsi le législateur a-t-il vu que la forme compte : une copie qui reproduirait le contenu sans la forme ne bénéficierait pas de la présomption.

Mais ce que ces textes protègent d'abord, c'est l'intégrité : la certitude que rien n'a changé. Ils ne disent pas que ce qui n'a pas changé pourra encore être lu. L'intégrité est garantie par des empreintes et des procédés qui certifient un objet, non par la capacité d'un logiciel à le restituer. Or la valeur d'une preuve suppose, en pratique, qu'on puisse en prendre connaissance. Un contrat intact, dont le fichier ne s'ouvre plus dans aucun logiciel disponible, est une preuve dont on peut établir qu'elle n'a pas été altérée, sans pouvoir établir ce qu'elle dit. La durabilité du contenant est ainsi la condition cachée de la valeur juridique. Aucun texte ne la formule en ces termes ; rien de ce que ces textes promettent ne tient pourtant sans elle.

La signature elle-même vieillit. On la vérifie au moyen de certificats qui expirent et de procédés de chiffrement qui cessent, avec les années, d'être tenus pour sûrs. Une signature

valable le jour de sa création peut ainsi devenir invérifiable si rien n'a été fait pour en conserver les preuves. La question n'a rien de théorique : une signature devenue invérifiable ne prouve plus rien. Le droit européen l'a prévu : il organise depuis 2014 un service de conservation des signatures électroniques qualifiées, destiné à en prolonger la valeur. C'est un remède curatif, apporté après coup à un défaut qui tient à la conception même de la signature électronique. Il a donc vu qu'une preuve non entretenue devient caduque. Il l'a vu pour la signature, qui certifie le contenu, avant de le voir pour le contenant qui le porte.

Il a fini par le voir. Le règlement européen sur l'identification électronique et les services de confiance, dit eIDAS, révisé en 2024, a introduit un nouveau service : l'archivage électronique. Il le définit comme un service qui reçoit, stocke, restitue et supprime des documents électroniques « afin d'en garantir la durabilité et la lisibilité », tout en préservant leur intégrité, leur confidentialité et la preuve de leur origine. Pour la première fois, un texte applicable en France fait de la lisibilité l'objet d'une garantie juridique. Une version qualifiée de ce service est prévue, dont un règlement d'exécution a depuis précisé les exigences techniques.

Cette avancée nuance la loi du visible : le droit nomme la lisibilité et l'encadre. Mais il la saisit à un endroit précis : dans un service que l'on choisit de confier à un prestataire, au moment où le document entre en archive. Il ne la saisit ni au moment où le document est produit, ni dans le logiciel qui le produit, ni dans le réglage par défaut qui décide de son format. La lisibilité est protégée à l'entrée du port d'arrivée. Elle ne l'est pas au chantier naval, au départ.

5.2 La dépossession de fait

Le droit du patrimoine offre aux archives publiques la protection la plus forte que connaisse le droit français. Le Code du patrimoine définit les archives comme l'ensemble des documents, y compris les données, produits ou reçus par toute personne ou tout organisme dans l'exercice de son activité. Il précise que ces documents en relèvent « quels que soient leur date, leur lieu de conservation, leur forme et leur support ». Il dispose que leur conservation est organisée dans l'intérêt public, tant pour la justification des droits des personnes que pour la recherche historique. Il pose enfin que les archives publiques sont imprescriptibles et que nul ne peut les détenir sans droit ni titre. Elles appartiennent en outre au domaine public, dont les biens ne peuvent être ni vendus ni acquis par le seul écoulement du temps.

Ces textes disent beaucoup. Une archive publique ne peut être ni vendue, ni appropriée, ni perdue avec le temps, et sa conservation sert la justification des droits comme la recherche historique, ce qui rejoint l'enjeu d'Histoire déjà évoqué. Mais tout cela porte sur

la propriété et la possession. L'imprescriptibilité protège le titre de l'État sur ses archives contre quiconque voudrait s'en emparer. Elle ne protège pas la capacité de l'État à les lire.

La définition elle-même le révèle. Les archives le sont « quels que soient [...] leur forme et leur support ». La formule est une garantie : aucun document n'échappe au régime parce qu'il serait numérique, parce qu'il serait un tableau plutôt qu'un texte, ou parce qu'il serait conservé dans un format plutôt que dans un autre. Mais cette neutralité a son revers : le régime protège toutes les formes sans en exiger aucune. Un fichier illisible est une archive publique au même titre qu'un registre relié, imprescriptible comme lui, inaliénable comme lui, et pourtant muet.

On peut ainsi conserver la propriété d'un fonds sans conserver la capacité de l'interroger. L'État demeure propriétaire, pour toujours, de ce qu'il ne pourra peut-être plus ouvrir. Et la loi du visible joue jusque dans ce régime. En effet, lorsqu'une administration confie ses archives courantes à un prestataire, le droit exige que celui-ci soit agréé et, pour le numérique, certifié selon la norme française des systèmes d'archivage électronique. Ce que le droit contrôle est donc le dépositaire et son système, non le format des fichiers qui y entrent : une fois encore, le port d'attache.

5.3 L'archivage, la douane du port d'arrivée

L'État n'est pourtant pas démuni. Il dispose d'une douane compétente.

Le programme interministériel d'archivage numérique de l'État, baptisé Vitam, a été lancé en 2015 par les ministères des Armées, de la Culture et des Affaires étrangères. Il a produit un logiciel libre d'archivage, déployé notamment aux Archives nationales. Ce logiciel sait identifier le format réel d'un fichier, quelle que soit l'extension sous laquelle il se présente. Il sait aussi vérifier qu'un fichier respecte bien la norme qu'il revendique, par exemple qu'un document qui se déclare conforme à la norme d'archivage du PDF l'est réellement. Des normes françaises encadrent ces pratiques, pour l'archivage électronique comme pour la numérisation fidèle des documents papier. C'est ici, et ici seulement, que la vérification de la déclaration de marchandise produit un effet. L'extension d'un fichier, le format annoncé par une ressource, la mention de conformité d'un PDF : tout ce qui voyageait sans contrôle est enfin confronté à la réalité.

Pourtant, cette douane a trois limites, qui tiennent à sa nature même. La première est son moment. Elle intervient au port d'arrivée, lorsque le document entre en archive, parfois des années après sa production. L'arrivée ne restitue pas ce qui n'a jamais été consigné à l'embarquement : l'encodage d'un fichier qui ne le déclarait pas doit être deviné, et une information que le logiciel a simplifiée à l'export n'existe plus. Ainsi l'archiviste peut identifier la coque ; il ne peut pas refaire le chargement.

La deuxième limite est son caractère facultatif. La documentation du programme présente la vérification des formats comme un service que le logiciel permet de mettre en œuvre, pour les administrations qui le souhaitent. L'outil existe ; la décision de s'en servir appartient à chaque administration.

La troisième limite est la plus inattendue ; la douane la reconnaît elle-même. Le registre des formats sur lequel elle s'appuie, PRONOM, pour identifier les fichiers est tenu par les Archives nationales britanniques, et plusieurs des logiciels de vérification qu'elle emploie viennent de l'étranger. Aussi la documentation du programme invite-t-elle les administrations à développer leur propre expertise, afin de ne pas être uniquement dépendantes d'expertises et d'outils étrangers. Il n'y a là aucun grief : ce registre est public, ouvert et tenu par les archives d'un État ami ; sa reprise serait possible. C'est une dépendance réversible, au sens où on l'a définie. Sans qu'il y ait là de reproche, le code du logiciel d'archivage de l'État est lui-même publié sur GitHub, la forge de Microsoft. Le seul contrôle suivi d'effet de toute la chaîne repose ainsi, lui aussi, sur une carte dressée ailleurs ; l'administration le reconnaît elle-même.

5.4 Le coffre et la clef

Tout ce qui précède vaut pour des documents que l'on sait lire. Le cas de la collection de textes préparée par l'État y ajoute une difficulté d'une autre nature. Une collection de textes préparée pour la recherche par le sens comprend deux choses : les textes eux-mêmes et les coordonnées produites par un modèle pour permettre d'y chercher. Les textes se conservent comme n'importe quelle archive. Mais les coordonnées ne s'interrogent qu'avec le modèle qui les a produites, et ce modèle est une pièce distincte, qui a son propre auteur, sa propre licence et sa propre durée de vie.

Qu'on archive les textes et leurs coordonnées sans archiver le modèle, et l'on conserve la propriété d'un ensemble dont on a perdu la clef. Les coordonnées sont intactes, imprescriptibles si elles sont publiques, et inutilisables. Tout devra alors être recalculé avec un autre modèle, à un coût que l'État n'a pas chiffré, pour des résultats qui ne seront pas ceux d'hier. La même question, posée dans un autre système de coordonnées, ne trouvera pas les mêmes réponses.

Rien n'empêche, techniquement, d'archiver la clef avec le coffre. Un modèle publié en libre accès peut être conservé à côté des textes qu'il a servi à classer. Encore doit-on avoir compris, pour le faire, que le modèle est une clef plutôt qu'un simple outil de passage et que sa conservation relève de la durabilité du fonds plutôt que de l'équipement informatique du moment. C'est une décision, du genre de celles que l'on a décrites comme abandonnées faute de consigne.

La durabilité retrouve ainsi sa place d'horizon. C'est à son aune que la chaîne entière doit être jugée, étage par étage. Le dernier venu des étages, celui du modèle, est celui où la question se pose avec le plus d'acuité et reçoit le moins de réponses.

Conclusion

Cet article ne désigne pas de bon format, ni de bon logiciel, ni de bon modèle. Il n'en existe pas en soi, seulement des navires adaptés à une traversée. En revanche, ce qu'il peut offrir au dirigeant est une méthode : une manière de poser la question du contenant avant que la cargaison ne parte. Et, pour ce qui doit traverser le plus long des voyages, il peut lui rappeler qu'il existe un contenant plus ancien que tous les autres.

La méthode tient en trois séries de questions, à poser à chacun des sept étages de la chaîne. La première est un plancher non négociable : la règle de ce contenant est-elle publique et complète, et échappe-t-elle au contrôle d'un seul acteur ? La deuxième porte sur la destination : ce document doit-il vivre, et donc rester modifiable, faire preuve, et donc rester figé, ou traverser le temps, et donc rester lisible dans vingt, cinquante ou cent ans ? La troisième est stratégique : ce contenant s'échange-t-il réellement sans perte ? Chaque étage passe enfin la même épreuve, celle de la réversibilité : qui a décidé du réglage par défaut, et que coûterait la sortie, en argent et en mois ?

Appliquée au fichier des subventions, la méthode aurait tenu en une ligne de consigne : publier dans l'encodage que la norme impose et le vérifier avant d'envoyer. Appliquée à la collection de textes préparée par l'État, elle aurait conduit à chiffrer, au moment de choisir le modèle, ce que coûterait un changement et à archiver la clef avec le coffre. Dans les deux cas, poser la question au bon moment aurait suffi, sans qu'il fût besoin d'en savoir davantage.

La nationalité d'un outil reste, on l'a vu, un indicateur : une dépendance étrangère réversible peut être moins grave qu'une dépendance proche dont on ne sait pas sortir. Le critère est la réversibilité, et son corollaire, vérifié sur le cas de l'État, est simple : une dépendance dont personne n'a chiffré la sortie n'a pas été choisie. Elle a été acceptée.

Cette méthode rend au dirigeant le geste qui lui appartient : fixer le cap. On l'a vu, il n'a pas à choisir lui-même chaque format ; il doit déléguer ce choix au lieu de l'abandonner, en donnant à ceux qui décident les critères de l'organisation plutôt que les leurs. Cet effort est modeste à l'échelle d'une décision. Il devient pourtant considérable lorsqu'on l'a différé pendant des années.

Une dernière question doit être posée, que le numérique n'aime pas qu'on lui pose. Parmi tous les contenants, il en est un dont la navigabilité ne dépend d'aucun voyage : le papier. Un document imprimé sur un papier de qualité ne demande ni logiciel, ni mise à jour, ni abonnement, ni électricité. Il ne tombe pas en panne, ne change pas de format, ne dépend d'aucune plateforme. Il reste lisible pour des générations qui ne partageront rien de notre environnement technique et constitue, depuis l'invention de l'imprimerie, la

base documentaire de l'Histoire. Sa conservation, on l'a vu, ne demande qu'un abri.

Le numérique a apporté des bienfaits considérables et incontestables : l'accès pour tous, la diffusion instantanée, la recherche, le partage des connaissances. La première phase de la transformation numérique a consisté à tout numériser : archives, procédures, actes, supports d'enseignement. Le mouvement a même atteint le cœur de l'État de droit : depuis 2016, le Journal officiel, qui publie les lois de la République, n'est plus imprimé ; il n'existe plus qu'en version électronique. Il est peut-être temps d'entrer dans une seconde phase, plus sélective : rematérialiser ce qui est essentiel, pour le sécuriser et le transmettre.

Revenir au tout papier n'aurait aucun sens. L'enjeu est, pour chaque catégorie de documents, de poser jusqu'au bout la question de la destination. Pour ce qui doit vivre, le numérique est irremplaçable. Mais pour ce qui doit traverser un siècle, la réponse la plus sûre est peut-être la matière.

L'achat d'un logement est, pour beaucoup de Français, l'acte patrimonial le plus important d'une vie. Il donne désormais lieu, dans la très grande majorité des cas, à un acte authentique électronique : le minutier central du notariat en comptait 35 millions au printemps 2025. Or l'acquéreur n'en reçoit souvent qu'un fichier. On pourrait concevoir qu'il en reçoive aussi un exemplaire imprimé sur un papier sécurisé, numéroté, vérifiable en ligne, destiné à être conservé et transmis. Le numérique porterait l'accès et la circulation ; l'imprimé porterait la preuve et la mémoire.

La vie économique en offre un autre exemple. Les comptes des entreprises, les liasses fiscales, les pièces justificatives sont aujourd'hui presque entièrement numériques. C'est efficace à court terme. À long terme, en revanche, une cyberattaque, la faillite d'un hébergeur, l'abandon d'un format ou une migration mal conduite peuvent rendre illisibles des pans entiers de la mémoire d'une entreprise. On peut imaginer qu'une synthèse essentielle, établie chaque année selon un modèle uniforme, soit conservée sous forme imprimée, hors d'atteinte des risques numériques : une capsule de mémoire, pour l'entreprise et, au-delà, pour le pays.

Les documents publics fournissent un dernier exemple. Décisions de justice, actes d'état civil, titres, autorisations : pour ceux qui fondent des droits sur la longue durée, la version numérique ne devrait peut-être pas être l'unique support. La matérialité joue ici un double rôle. Elle protège le document des aléas de la donnée. Elle dit aussi, symboliquement, le poids que la puissance publique attache à ce qu'elle décide. La création d'une entreprise en offre une illustration. Dans un univers où elle n'est plus qu'une suite de formalités en ligne, un acte de fondation imprimé, authentifié et digne d'être affiché rappellerait que l'on vient de créer une institution et d'en prendre la responsabilité.

Mais la rematérialisation a ses contre-indications ; la vie des entreprises en fournit la

meilleure illustration. Un extrait du registre du commerce, imprimé ou figé dans un PDF, atteste de la situation d'une société au jour où il a été délivré. Pourtant, il continue de circuler bien après, alors que la société a pu changer de dirigeant, déménager ou perdre toute activité réelle. Un document figé devient alors l'instrument de la tromperie : il prouve ce qui n'est plus vrai. Pour une information qui doit refléter un état vivant, la réponse est un registre numérique tenu à jour, authentifié par un tiers de confiance et consulté au moment où l'on en a besoin. On retrouve la question de la destination : la matière pour ce qui ne doit plus changer, le numérique authentifié et vivant pour ce qui doit rester exact. Le papier n'est pas la réponse à tout ; il est la réponse à ce qui doit durer tel quel.

Le cinéma, déjà évoqué, en donne la mesure : c'est la pellicule, support matériel, qui a permis de restaurer les œuvres anciennes, là où les supports plus modernes qui l'avaient remplacée ont fixé pour toujours leur médiocrité. Cette voie a toutefois ses exigences. Le papier n'est pas éternel par nature : les papiers acides des XIX^e et XX^e siècles se désagrègent, et il existe des normes conçues pour le papier destiné à durer. Imprimer n'est pas davantage conserver, si l'exemplaire imprimé n'est pas lui-même archivé avec soin. Mais la rematérialisation raisonnée a un mérite que n'a aucun format numérique : elle rend une partie de la mémoire indépendante de tout outil. On dit souvent que l'avenir est au tout numérique. Il est plus probablement hybride : le numérique pour le flux, l'imprimé pour ce qui doit rester. Imprimer ce qui compte est un acte de responsabilité envers ceux qui viendront après nous et qui chercheront à comprendre ce que nous avons décidé ; la nostalgie n'y a aucune part.

On dit volontiers que le numérique a rendu le papier inutile. C'est l'inverse qui est vrai : en confiant l'essentiel de notre mémoire à des contenants qui ne durent que tant qu'on les entretient, le numérique a rendu le papier indispensable.

Cette revue n'échappe pas à ce qu'elle décrit ; il serait malhonnête de ne pas l'écrire. Elle n'existe que par le numérique, au moyen de contenants qu'elle n'a pas tous choisis et dont aucun ne lui garantit vingt ans de lisibilité. Que restera-t-il de ces pages dans cinq ans, dans dix ans, si personne ne renouvelle un nom de domaine, ne migre un format, ne conserve ailleurs ce qui aura été écrit ? La question ne se pose pas autrement pour une revue que pour un ministère ; c'est même ce qui la rend sérieuse. Un texte qui examinerait les contenants des autres sans regarder le sien resterait, au sens propre, un écrit hors sol.

Reste le droit. Cet article n'a pas soutenu que la norme se désintéresse volontairement de la coque et de la carte. Il a soutenu qu'elle règle ce qu'elle voit et qu'elle ne voit que le pavillon. Pourtant, cette explication a une durée. Un angle mort cesse d'en être un le jour où il a été signalé. Les bas-fonds sont désormais décrits : les faits sont mesurés, les textes sont cités, et un État voisin a même inscrit la carte parmi les normes qui lui manquent.

Continuer de ne régler que le pavillon ne relèverait plus du retard du droit sur son objet. Ce serait un choix. La revue ne dit pas qu'il y a hypocrisie; elle retire seulement l'excuse de l'ignorance.

Le dernier mot revient au geste le plus humble de toute la chaîne, celui de l'introduction. Quelque part, un agent a cliqué sur « Enregistrer », et son logiciel a choisi pour lui. Rien n'est plus ordinaire. C'est pourtant à cet instant, et non dans les grandes lois, que le logiciel a décidé de ce qu'un lecteur pourra lire, dans dix ans, d'une décision de l'État. La souveraineté se proclame au pavillon. Elle se joue dans la coque.

Sources et références

Introduction

Arrêté du 17 novembre 2017 relatif aux conditions de mises à disposition des données essentielles des conventions de subvention, pris pour l'application du décret n° 2017-779 du 5 mai 2017; annexe (« Le jeu de caractère (encoding) à utiliser est UTF-8 »), version en vigueur depuis le 20 novembre 2017, consultée sur Légifrance.

Mesures de la revue du 19 septembre 2026 sur data.gouv.fr. Jeu de données des subventions accordées par le cabinet du Premier ministre de 2017 à 2021, publié par l'organisation certifiée « Premier ministre » et déclaré conforme au schéma national « Subventions » : cinq fichiers annuels hors UTF-8, en-tête « Référence de la décision » défiguré, absence de licence, plus de 8 000 téléchargements, dernier contrôle du robot de la plateforme en février 2026 (« texte, sans erreur »). Échantillon de 144 fichiers CSV publics publiés de 2016 à 2025, dont 9 hors UTF-8.

Exemples du quotidien. Apple, *Using HEIF or HEVC media on Apple devices*, documentation d'assistance : formats HEIF et HEVC introduits avec iOS 11 et recommandés par défaut, réglage Appareil photo, Formats, permettant de revenir aux formats « les plus compatibles ». Affichage « Ã© » d'un « é » encodé en UTF-8 et lu selon une table Latin-1. Scène des présentations préparées sous PowerPoint et projetées sous suite bureautique libre dans une administration : témoignage de l'auteur.

La doctrine du contenant numérique

Droit. Code civil, articles 1366, 1367 et 1379, dans leur rédaction issue de l'ordonnance n° 2016-131 du 10 février 2016. Convention internationale pour l'unification de certaines règles en matière de connaissance, Bruxelles, 25 août 1924, article 3, paragraphes 1 et 2; sur la portée de la formule « avant et au début du voyage », Comité judiciaire du Conseil privé, *Maxine Footwear Co. Ltd v. Canadian Government Merchant Marine Ltd*, [1959] A.C. 589, [1959] 2 Lloyd's Rep. 105. Code des transports, article L. 5422-6, issu de l'ordonnance n° 2010-1307 du 28 octobre 2010, consulté sur Légifrance.

Disparition et archivage du web. Pew Research Center, *When Online Content Disappears*, 17 mai 2024 (38 % des pages existant en 2013 inaccessibles en octobre 2023). Scott G. Ainsworth, Michael L. Nelson et Herbert Van de Sompel, « Only One Out of Five Archived Web Pages Existed as Presented », *Proceedings of the 26th ACM Conference on Hypertext and*

Social Media, 2015. Adobe, *Flash Player End of Life* (fin du support le 31 décembre 2020, blocage des contenus à partir du 12 janvier 2021); Internet Archive et émulateur libre Ruffle; Smithsonian Institution Archives, *Gone Flash*, sur la difficulté de capture de ces contenus par les outils d'archivage du web. Dépôt légal du web : loi n° 2006-961 du 1^{er} août 2006 et son décret d'application de 2011; Bibliothèque nationale de France, *Comprendre le dépôt légal du web* (collecte par robot, sans démarche du producteur; collectes larges annuelles et collectes ciblées).

Photographie. Formats bruts propres à chaque fabricant (RAW), format JPEG.

Gouvernance. Michael C. Jensen et William H. Meckling, « Theory of the Firm : Managerial Behavior, Agency Costs and Ownership Structure », *Journal of Financial Economics*, vol. 3, n° 4, octobre 1976, p. 305-360. Publications de la revue : N°02, *La gouvernance silencieuse*; hors-série n° 1, *La sûreté de l'intention*.

Les étages de la chaîne numérique

Donnée. Mark Ziemann, Yotam Eren et Assam El-Osta, « Gene name errors are widespread in the scientific literature », *Genome Biology*, vol. 17, 2016, article 177, DOI 10.1186/s13059-016-1044-7 (environ un cinquième des articles accompagnés de listes de gènes au format tableur affectés). Elspeth A. Bruford et alii, « Guidelines for human gene nomenclature », *Nature Genetics*, vol. 52, 2020, p. 754-758, et annonce du comité de nomenclature HGNC d'août 2020 (symboles modifiés pour éviter leur conversion automatique en dates : SEPT_I devenu SEPTIN_I, MARCH_I devenu MARCHF_I).

Formats. OpenDocument : standard OASIS du 1^{er} mai 2005, norme ISO/IEC 26300 du 30 novembre 2006. Office Open XML : Ecma, ECMA-376, 2006; norme ISO/IEC 29500, édition de 2008, variantes stricte et transitoire, dont l'objet déclaré est de représenter fidèlement le corpus existant des documents produits par Microsoft Office. Library of Congress, *Sustainability of Digital Formats*, fiche fdd000397, « DOCX Transitional (Office Open XML) » : tests de février 2017 montrant que Google Docs et LibreOffice créent leurs nouveaux documents dans la variante transitoire, comme Word par défaut depuis 2013. PDF : ISO 32000-1 du 1^{er} juillet 2008; PDF 2.0, ISO 32000-2, éditions de 2017 et de 2020, librement téléchargeable depuis avril 2023; PDF/A, ISO 19005-1, 2005. Markdown : John Gruber, avec Aaron Swartz, 2004; spécification CommonMark, 2014; RFC 7763 et 7764, 2016. JSON : RFC 8259, décembre 2017, section 8.1. CSV : RFC 4180, octobre 2005, document de statut informatif et non norme. Format .pages du traitement de texte d'Apple.

Encodage et fichiers publics. Constat de la revue du 20 septembre 2026 : une ressource du répertoire national des élus publiée sur data.gouv.fr est encodée en UTF-8 conformément

à la règle, mais dépourvue de la marque d'ordre des octets qui signale l'encodage ; ouverte dans un tableur réglé sur la table Windows-1252, elle affiche « dÃ@partement » dès l'entête ; séparateur : point-virgule. Mesure de la revue du 20 septembre 2026 : 39 ressources au format .doc publiées par les services du Premier ministre. Ressource de la ville de Béthune déclarée au format xls et écrite en SpreadsheetML 2003.

SMS. 3GPP TS 23.038 (alphabet GSM à 7 bits, 160 caractères ; UCS-2, 70 caractères) et TS 23.040, clause 9.2.3.24.1 (messages concaténés : 153 ou 67 caractères par segment) ; charge utile de 140 octets ; bascule de tout le message en UCS-2 dès qu'un caractère étranger à l'alphabet GSM est présent (émojis, apostrophes courbes, tirets longs).

Outils. Euro-Office, lancé le 9 juin 2026, construit sur le code d'OnlyOffice (licence AGPL), coexistant avec Collabora dans Nextcloud. Grist, Grist Labs (New York), adopté et hébergé par l'État. MapLibre GL JS, bifurcation de décembre 2020, licence BSD ; cartes.gouv.fr, ouvert le 25 juin 2026 par l'IGN.

Guidage par satellite. GPS.gov, *Selective Availability* : dégradation volontaire du signal civil, précision d'environ 100 mètres, supprimée le 1^{er} mai 2000 sur décision présidentielle. Système européen Galileo, services initiaux ouverts le 15 décembre 2016.

Modèle. Jianlv Chen, Shitao Xiao, Peitian Zhang, Kun Luo, Defu Lian et Zheng Liu, « M3-Embedding : Multi-Linguality, Multi-Functionality, Multi-Granularity Text Embeddings Through Self-Knowledge Distillation », arXiv, identifiant 2402.03216, première version du 5 février 2024 (Académie d'intelligence artificielle de Pékin et Université de science et technologie de Chine). Service hydrographique et océanographique de la Marine : Code de la défense, articles R. 3416-1 et suivants ; Dépôt des cartes et plans de la Marine, 1720.

Hébergement. Plateforme des données de santé, anciennement Health Data Hub, créée par la loi du 24 juillet 2019, hébergée sur Microsoft Azure ; annonce ministérielle du 6 février 2026 de la migration directe vers un opérateur français ; Assemblée nationale, rapport n° 3054, tome I, 8 juillet 2026, p. 202 à 207.

Canal. TeleGeography (plus de 99 % de la capacité intercontinentale portée par les câbles sous-marins) ; ministère des Armées, stratégie ministérielle de maîtrise des fonds marins, février 2022. Moteurs de recherche : Brian Dean, « We Analyzed 4 Million Google Search Results. Here's What We Learned About Organic Click Through Rate », Backlinko, consulté le 22 septembre 2026 (données de 1 312 881 pages et 12 166 560 requêtes ; 0,63 % des internautes cliquent sur un résultat de la deuxième page). Hugging Face : société fondée en 2016 à New York par Clément Delangue, Julien Chaumond et Thomas Wolf ; SAS Hugging Face, SIREN 822 168 043, présidée par Hugging Face, Inc. depuis une décision de l'associé unique du 20 juin 2025 publiée au BODACC ; NVIDIA, « NVIDIA to Acquire Hugging Face », blog de l'entreprise, 3 septembre 2026 (accord d'acquisition). GitHub,

propriété de Microsoft depuis 2018; OpenAnafi (Cour des comptes); pôle logiciels libres de la DINUM; SourceHut; recensement public des dépôts de code des administrations, code.gouv.fr. Référentiel INGRES du CISIRH, placé sous licence ouverte, dont l'interface est réservée au réseau interministériel de l'État.

Poste. Serveur expérimental de data.gouv.fr destiné aux assistants d'intelligence artificielle, interrogé le 5 septembre 2026 sur le répertoire national des élus : quatre écarts, dont trois issus des métadonnées du producteur.

Licence. Mesure de la revue du 19 septembre 2026 : 27,8 % des jeux publics non archivés de data.gouv.fr sans licence valide (« non spécifiée » ou vide). Open Data Commons, *Open Database License*, version 1.0, article 4.6. Lignes directrices de la fondation OpenStreetMap sur les jeux d'entraînement; discussion contradictoire sur la liste *legal-talk* d'OpenStreetMap, avril 2019, « OSM for training ML machines ».

Une loi de surface

Cyberattaque de la DGFIP. Communiqué du ministère du 14 août 2026 : accès frauduleux de juin et juillet 2026 par usurpation des identifiants d'un agent et d'un tiers habilité; 678 000 particuliers et professionnels concernés; atteinte à des outils internes, non à l'hébergement.

Hébergement. Circulaire du Premier ministre du 8 novembre 2018 sur la doctrine d'utilisation de l'informatique en nuage par l'État, rappelée dans l'historique de la circulaire n° 6282-SG du 5 juillet 2021 dite « cloud au centre » (NOR : PRMX2121516C), dont la règle [R9] dispose que, si le système manipule des données d'une sensibilité particulière, « l'offre de cloud commercial retenue devra impérativement respecter la qualification SecNumCloud (ou une qualification européenne d'un niveau au moins équivalent) et être immunisée contre toute réglementation extracommunautaire ». Circulaire n° 6404/SG du 31 mai 2023 (NOR : PRMG2314778C), qui délimite le périmètre des données d'une sensibilité particulière et encadre les dérogations. Loi n° 2024-449 du 21 mai 2024 visant à sécuriser et à réguler l'espace numérique, article 31. Décret n° 2026-272 du 14 avril 2026. Arrêté du 12 août 2026 portant approbation du référentiel d'exigences relatif aux prestataires de services d'informatique en nuage (NOR : PRMD2617964A), approuvant la version 3.2 du référentiel SecNumCloud, publié au *Journal officiel* du 14 août 2026. CNIL, délibérations n° 2025-013 et 2025-014 du 13 février 2025; Conseil d'État, 20 mars 2026, n° 503159 et 504171 (rejet des recours contre l'autorisation de la CNIL : données hébergées en France, absence de transfert de données de santé, garanties contractuelles pour les données des utilisateurs); CJUE, 16 juillet 2020, C-311/18, *Schrems II*.

Certification européenne. Audition de M. Henri d'Agrain, délégué général du Cigref, devant

la commission d'enquête le 30 avril 2026, rapportée au rapport n° 3054, p. 239-240 : niveau renforcé « High + » du schéma EUCS, intégrant des critères d'immunité aux législations extraterritoriales, jusqu'à la mi-2023; courrier des associations CCIA et BSA à l'administration américaine du 25 mai 2023; note diplomatique américaine de septembre 2023; schéma réapparu en mars 2024 sans ces critères.

Format. Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique, article 16. Référentiel général d'interopérabilité, version 2.0, arrêté du 20 avril 2016. Rapport n° 3054 précité, proposition n° 12 de la rapporteure, p. 390 et 415. Arrêté du 17 novembre 2017 précité. Mission Etalab, dépôt *format-subventions* (choix du point décimal « contrairement à la lettre de l'arrêté »); schéma *Subventions* sur schema.data.gouv.fr.

Comparaisons. IT-Planungsrat, décision 2025/06 du printemps 2025 (formats ouverts, dont ODF, appelés à devenir d'ici 2027 le standard de l'échange de documents); décision B-2025/44-IT (prise en charge progressive d'ODF par les administrations des Länder à compter du 1^{er} janvier 2027, évaluation en 2028); décision B-2026/03-IT, *Deutschland-Stack*, 18 mars 2026, et ses annexes *Portfolio* et *Standards*, consultées le 20 septembre 2026, dont la liste des besoins de normalisation (format d'intégration des sources pour la génération augmentée, format d'échange des modèles de langage, format de stockage des données vectorielles). Bavière : heise online, 12 janvier 2026 (désaccord public au sein du gouvernement du Land) et 3 juin 2026 (abandon des négociations); volume de près d'un milliard d'euros sur cinq ans rapporté par le magazine c't d'après une présentation du ministère bavarois des Finances. Schleswig-Holstein : communiqué du gouvernement du Land du 4 décembre 2025. Estonie : *Estonian IT Interoperability Framework*, version 3.0, décembre 2011 (liste de standards ouverts, dont ODF, PDF/A et CSV); nouveau cadre de 2023, document évolutif et non contraignant, selon la fiche *Digital Public Administration 2024* consacrée à l'Estonie par l'observatoire européen des cadres nationaux d'interopérabilité (NIFO). Norvège : *Referanse katalog for IT-standarder i offentlig sektor*, version 2.0, 25 juin 2009; règlement du 5 avril 2013 n° 959 sur les standards informatiques dans l'administration publique.

Union européenne. Règlement (UE) 2024/903 du 13 mars 2024 pour une Europe interopérable, article 3. Directive 2014/24/UE sur les marchés publics, article 42, paragraphe 4 (interdiction de la référence à une marque, un brevet, une origine ou une production déterminée, sauf justification par l'objet du marché, avec la mention « ou équivalent »). Règlement d'exécution (UE) 2023/138 de la Commission du 21 décembre 2022 sur les ensembles de données de forte valeur (licences CCo, CC BY 4.0 ou équivalente ou moins restrictive), consulté sur EUR-Lex; la catégorie « observation de la Terre et environnement » de son annexe couvre les informations environnementales au sens de la directive 2003/4/CE et les données produites en application des actes qu'elle énumère, dont la lé-

gislation sur la qualité de l'air ambiant (directive 2008/50/CE, remplacée par la directive (UE) 2024/2881). Règlement (UE) 2024/1689 sur l'intelligence artificielle, articles 1^{er} et 2 (objet et champ d'application).

Changement de fournisseur. Règlement (UE) 2023/2854 du Parlement européen et du Conseil du 13 décembre 2023 concernant des règles harmonisées portant sur l'équité de l'accès aux données et de l'utilisation des données (règlement sur les données), *Journal officiel de l'Union européenne*, série L, 22 décembre 2023, chapitre VI, notamment articles 23, 26, 29 et 30, ce dernier en ses paragraphes 5 (format d'export, à défaut de norme commune publiée) et 6 (actifs protégés par la propriété intellectuelle ou le secret des affaires); applicable depuis le 12 septembre 2025. Règlement (UE) 2016/679 (RGPD), article 20. Loi n° 2024-449 du 21 mai 2024 précitée, article 27. ARCEP, décision n° 2025-0340 du 20 février 2025. Arrêté du 17 novembre 2025 fixant le montant maximal de tarification pour les frais de transfert de données dans le cadre d'un changement de fournisseur de services d'informatique en nuage (montant fixé à zéro euro), consulté sur Légifrance. Commission européenne, proposition de règlement modifiant notamment le règlement (UE) 2023/2854 (« omnibus numérique »), COM(2025) 837 final, 19 novembre 2025, en cours d'examen par le Parlement européen et le Conseil; le texte de compromis du Conseil exempte des obligations de changement de fournisseur, pour les contrats conclus avant le 12 septembre 2025, les services personnalisés et ceux des PME et des petites entreprises de taille intermédiaire, hors fourniture d'infrastructure, le régime des frais de l'article 29 restant applicable. Pour une lecture convergente du chapitre VI, appliquée aux données de gestion d'actifs : Nizar Younes Mqam, « La donnée d'actif : la sortie devient gratuite, le modèle reste captif », *Cahiers de technopolitique*, n° 3, 18 septembre 2026, DOI 10.5281/zenodo.22836960.

Qualité de l'air. Atmo France, *FAQ API Atmo Data*, 10 octobre 2023 : données des associations agréées de surveillance de la qualité de l'air diffusées sous licence ODbL; jeux de ces associations publiés sur data.gouv.fr sous la même licence.

Codes sources. Code des relations entre le public et l'administration, article L. 300-2. Circulaire du Premier ministre du 27 avril 2021 sur la politique publique de la donnée, des algorithmes et des codes sources. Rapport n° 3054 précité, recommandations n° 17 et 18.

Mesure de la plateforme. Mesure de la revue du 19 septembre 2026 sur 393 355 ressources vivantes de data.gouv.fr : 2 132 valeurs distinctes du format déclaré; concordance de 96,6 % entre format déclaré et format réel sur le périmètre comparable, qui couvre 37 % des ressources; 0,5 % d'écarts pour les fichiers hébergés par la plateforme; un quart des ressources jamais analysé; 7,4 % inaccessibles; 9 fichiers CSV hors UTF-8 sur 144 dans l'échantillon. Lecture du code publié de la plateforme (robot d'analyse et interface).

Rapport parlementaire. Assemblée nationale, rapport n° 3054 fait au nom de la commission d'enquête sur les dépendances structurelles et les vulnérabilités systémiques dans le secteur du numérique et les risques pour l'indépendance de la France, président M. Philippe Latombe, rapporteure Mme Cyrielle Chatelain, tome I, 8 juillet 2026, 453 pages; 45 auditions et 112 personnes entendues (compte rendu n° 47, séance du 8 juillet 2026); dix-huit recommandations de la commission et propositions de la rapporteure numérotées de 1 à 30; audition favorable à ODF, p. 365-366; recherche en texte intégral : aucune occurrence des termes *embedding*, vectorisation, encodage ni UTF-8.

Sous-section 3.9. Mesure de la revue du 20 septembre 2026 sur data.gouv.fr, jeux non archivés de cinq organisations certifiées; échantillon aléatoire de douze fichiers CSV au plus par organisation, test de décodage UTF-8 sur les 200 premiers kilo-octets. Premier ministre : 179 jeux, 535 ressources, 13 jeux à licence non spécifiée, 219 PDF, 39 .doc, 10 CSV non UTF-8 sur 12 (1 inaccessible). Assemblée nationale : 11 jeux, 23 ressources, formats JSON et XML dominants, 2 CSV testés non UTF-8. Sénat : 8 jeux, 86 ressources dont 79 hébergées hors plateforme, 10 CSV non UTF-8 sur 12, encodage Windows-1252 constaté (« Requête »). Conseil constitutionnel : 4 jeux, 49 ressources dont 47 sur la plateforme, 8 CSV UTF-8, 4 inaccessibles. Insee : 198 jeux, 194 sous licence ouverte, 517 ressources dont 194 sans format déclaré, 43 au format Parquet, 11 CSV UTF-8, 1 inaccessible. Cour des comptes non mesurée, faute de réponse stable de la plateforme.

Le régulateur face à ses choix

Collection MediaTech. Compte AgentPublic sur la plateforme Hugging Face, neuf jeux de données consultés le 20 septembre 2026; modèle BAAI/bge-m3 indiqué dans chaque fiche; qualification *semantic-ready* dans les fiches de l'annuaire des administrations de l'État et du catalogue de data.gouv.fr; code source *etalab-ia/mediatech*; fiche beta.gouv.fr du projet (diffusion sur Hugging Face et sur data.gouv.fr). Jeu LEGI : 2 195 690 passages vectorisés au 20 septembre 2026.

Modèle. Fiche du modèle BGE-M3 sur Hugging Face, huggingface.co/BAAI/bge-m3 : licence MIT, prolongement de XLM-RoBERTa (Facebook AI), téléchargements mensuels relevés en septembre 2026. Publication du modèle : Chen et alii, arXiv 2402.03216, précitée. Beijing Academy of Artificial Intelligence, www.baai.ac.cn/en. Bureau of Industry and Security, *Additions to the Entity List, Federal Register*, vol. 90, n° 59, 28 mars 2025, document 2025-05427, effectif au 25 mars 2025. Bancs d'essai publics : Niklas Muennighoff, Nouamane Tazi, Loïc Magne et Nils Reimers, « MTEB : Massive Text Embedding Benchmark », arXiv, identifiant 2210.07316, 2022.

Albert. Catalogue d'Albert API constaté le 19 septembre 2026 (*bge-m3* et *bge-reranker-v2-m3*); rapport n° 3054 précité, p. 70 (« modèles d'IA open source et français »). Robertson

Davies, *Tempest-Tost*, Toronto, Clarke, Irwin & Co., 1951 (« The eye sees only what the mind is prepared to comprehend », traduction de l'auteur) ; l'attribution à Henri Bergson, fréquente, n'est pas étayée (Quote Investigator, 8 février 2018).

Commission européenne. Contrôleur européen de la protection des données, affaire 2021-0518 : décision du 8 mars 2024, lettre de clôture du 11 juillet 2025, communiqué du 28 juillet 2025. Bureau des programmes open source de la Commission (OSPO), créé en 2020 au titre de la stratégie open source 2020-2023 ; éditeur législatif libre LEOS (*Legislation Editing Open Software*), publié sous licence EUPL.

La pérennité du contenant

Code civil, articles 1366 et 1379 ; décret n° 2016-1673 du 5 décembre 2016 relatif à la fiabilité des copies, articles 2 à 6. Règlement (UE) n° 910/2014 dit eIDAS, article 34 (service qualifié de conservation des signatures électroniques qualifiées), ainsi que l'article 3, point 48, et l'article 45 *undecies* relatifs à l'archivage électronique, dans leur rédaction issue du règlement (UE) 2024/1183 du 11 avril 2024 (texte consolidé consulté sur EUR-Lex) ; règlement d'exécution (UE) 2025/2532 du 16 décembre 2025. Code du patrimoine, articles L. 211-1, L. 211-2 et L. 212-1 ; Code général de la propriété des personnes publiques, articles L. 2112-1, 2°, et L. 3111-1.

Programme Vitam, lancé le 9 mars 2015 ; identification par l'outil Siegfried sur le registre PRONOM des Archives nationales du Royaume-Uni ; validation par des outils tels que JHOVE et veraPDF ; documentation fonctionnelle, *Validation du format des fichiers*. Normes NF Z 42-013, version d'octobre 2020, et NF Z 42-026, mai 2017. Code source du logiciel Vitam publié sur GitHub, dépôt ProgrammeVitam/vitam, constaté le 20 septembre 2026.

Conclusion

Loi organique et loi du 22 décembre 2015 portant dématérialisation du *Journal officiel* de la République française, publication exclusivement électronique à compter du 1^{er} janvier 2016 ; Conseil constitutionnel, décision n° 2015-724 DC du 17 décembre 2015 (faculté pour toute personne d'obtenir sur papier l'extrait d'un acte publié). Groupe ADSN, communiqué du 27 mars 2025 : trente-cinq millions d'actes authentiques électroniques déposés au minutier central électronique du notariat. Norme ISO 9706, édition de 1994, *Information et documentation : papier pour documents, prescriptions pour la permanence*. Extrait du registre du commerce et des sociétés, figé à sa date de délivrance ; registre tenu par les greffiers des tribunaux de commerce, officiers publics et ministériels.

L'auteur



Sébastien VANNEROT

Juriste de formation et fin connaisseur des rouages de la décision collective, l'auteur a dirigé pendant cinq ans une organisation patronale de premier plan. Cette expérience de terrain, alliée à une culture de la rigueur et du résultat, nourrit une réflexion constante sur la sécurité économique et la prospective stratégique. Co-auteur de référentiels nationaux sur la sûreté des organisations, il s'attache à transformer la complexité réglementaire en leviers de résilience opérationnelle. Sa plume, à la fois ferme et bienveillante, explore les sous-jacents qui font dérailler les systèmes, des dérives du juridisme aux impasses de la gouvernance silencieuse. À travers ses écrits sur l'économie du réel, la souveraineté productive et la refondation d'un État responsable, il plaide pour un retour aux fondamentaux physiocratiques et ponocratiques. Son approche, résolument constructive, cherche à identifier les failles pour mieux bâtir des architectures de confiance, en ne perdant jamais de vue l'aspect social et humain de l'entreprise. À la croisée du droit, de la fiscalité et de la stratégie, il propose une vision de la gouvernance où la parole libérée et la lucidité sont les conditions nécessaires à la pérennité des structures complexes.

PARADOXALE

LE LAB DES CONTRE-SENS

Direction de la publication et de la rédaction :

Sébastien VANNEROT (*personne physique éditrice*)

Rédaction et analyses : Le Lab des contre-sens

Contact : contact@paradoxale-larevue.fr

Site : <https://paradoxale-larevue.fr>

PARADOXALE[®] est une marque enregistrée à l'INPI sous le numéro 26 5255335.

Date de publication : octobre 2026

N° ISSN : 2981-2771 (en ligne)

Cette revue, publiée sous forme numérique, est soumise au dépôt légal numérique de la Bibliothèque nationale de France conformément aux articles L131-1 et suivants du Code du patrimoine. Le contenu est moissonné par les robots de la BnF au titre du dépôt légal de l'internet.

Les citations issues des articles et travaux de Paradoxale sont bienvenues, sous réserve d'indiquer clairement la source : « Paradoxale – Le Lab des contre-sens, N° 10, octobre 2026, Sébastien VANNEROT »..

Ce numéro est diffusé sous licence Creative Commons Attribution – Pas d'Utilisation Commerciale – Pas de Modification 4.0 International (CC BY-NC-ND 4.0). Sa reproduction et sa diffusion intégrales sont libres et gratuites, sous réserve de citer la source dans la forme indiquée ci-dessus, à l'exclusion de tout usage commercial et de toute modification. Outre les droits accordés par cette licence, l'auteur autorise la diffusion interne au sein des entreprises, administrations et établissements d'enseignement, y compris à but lucratif. La reproduction dans un support éditorial diffusé à titre onéreux demeure soumise à autorisation préalable. Le droit moral de l'auteur demeure inaliénable.

La présente licence porte sur les seuls contenus éditoriaux. Elle ne s'étend pas aux signes distinctifs de la publication, dont la marque verbale et le logotype demeurent la propriété exclusive de leur titulaire.

PARADOXALE

POURQUOI LES MEILLEURES DÉCISIONS MÈNENT PARFOIS À L'ÉCHEC ?

Pourquoi notre société et nos organisations produisent-elles si régulièrement des résultats inverses à leurs intentions initiales ? Parfois parce qu'une décision fut franchement mauvaise et il en existe. Mais bien plus souvent parce que, prise isolément, chaque décision se défend. Aucune n'est totalement absurde, aucune n'a réellement de coupable. C'est leur accumulation qui finit par produire un résultat que personne n'a souhaité et dont personne n'assume les conséquences. Rien n'est plus difficile à défaire qu'un ensemble dont chaque pièce répond défensivement à une autre.

Chaque échelon décide dans son périmètre mais la cohérence de l'ensemble n'appartient à personne et aucun regard extérieur ne vient réellement l'éprouver. La norme prospère dans cet angle mort. À force de tout vouloir prévoir, jusqu'au plus petit détail, nous avons fait du droit une finalité alors qu'il n'est qu'un moyen, et de la conformité un alibi qui dispense de penser et de juger. Une loi inutile affaiblit les lois nécessaires et pourtant, de l'inverse, nous avons fait un choix de société.

La revue **Paradoxale** est l'espace dédié au décryptage de ces trajectoires dévoyées. Chaque numéro explore les points de bascule où le consensus se transforme en inertie, où la norme sert des intérêts particuliers au nom de l'intérêt général, où la règle produit l'inverse de l'effet économique attendu et où le silence des acteurs précipite l'erreur collective. Paradoxale n'appartient à aucun camp et ne sert aucune cause partisane, elle tente seulement, par ses analyses, de contribuer au débat public et à la poursuite de l'intérêt général.

En croisant la précision du raisonnement juridique, l'analyse des mécanismes économiques et la réalité des rapports humains, Paradoxale s'attache à mettre en lumière les sous-jacents qui font dérailler les stratégies les plus abouties et qui fragilisent, sans qu'on les voie venir, la sécurité économique et juridique des acteurs. Le regard y porte moins sur les décisions prises une à une que sur la manière dont elles s'accumulent et se contraignent les unes les autres. Elle invite les décideurs, publics comme privés, à regarder au-delà des évidences pour identifier les failles structurelles et les angles de rupture de nos cadres de pensée.

Traquer le sens derrière le contre-sens : telle est l'ambition de cette revue. Rendre à la décision sa pleine lucidité, non pour la dénoncer, mais pour l'expliquer, car le vrai danger n'est pas le désaccord, c'est l'accord que personne ne questionne.