

ANONYMISER / DÉSANONYMISER DES DONNÉES AVEC L'AIDE D'UNE IA GÉNÉRATIVE

Alain MICHEL
Agnès CONSTANT
Jérôme BOUSQUET-CARTON

AUTEURS :

- Alain MICHEL
- Agnès CONSTANT
- Jérôme BOUSQUET-CARTON

LICENCES :

Creative Commons - Attribution - Partage dans les Mêmes Conditions

AEFE – Bureau de la Formation et de l'Innovation – Nov. 2025

Table des matières

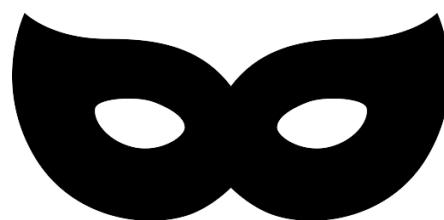
I. EXEMPLE D'UTILISATION.....	3
II. COMMENT FAIRE ?.....	4
III. RESSOURCES ASSOCIÉES.....	11
IV. RECOMMANDATIONS D'USAGE ET JURIDIQUES.....	12

Présentation

Les outils d'IA génératives actuels peuvent nous faire gagner un temps considérable dans certaines situations auxquelles on ne pense pas forcément au tout premier abord. Par exemple, si nous souhaitons travailler sur des données contenues dans un tableur, la tentation sera grande de les faire analyser par une IA pour en extraire des tendances, ou une synthèse, ou générer un traitement particulier.

Sauf qu'il est impensable de soumettre des fichiers de données structurées dans un tableur qui peuvent contenir des informations à caractère personnel à une IA, quelle qu'elle soit.

Il va donc être **indispensable d'anonymiser les données contenues dans le fichier à traiter AVANT de le passer à la moulinette de l'IA** pour en obtenir le traitement désiré. Une fois ce traitement effectué, on pourra de la même manière désanonymiser le fichier pour retrouver ses données complètes.



↓ [Télécharger une version PDF^{\[p.\]}](#) de ce tutoriel.

 Pour faire ce travail d'anonymisation des données contenues dans un fichier, nous allons utiliser une IA générative pour créer le code d'une page HTML qui effectuera localement ce travail à partir du fichier fourni.

I. Exemple d'utilisation

Créer un outil d'anonymisation de données avant leur traitement par une IA

Hypothèse de départ : vous êtes chef ou cheffe d'établissement et vous devez analyser des tableaux de données extraites de votre base de données établissement et contenant évidemment des données nominatives d'élèves et de parents. Pour faire l'analyse que vous souhaitez, vous auriez besoin de faire analyser les données par une IA générative.

Problème : ces fichiers contiennent des données sensibles, nominatives et très personnelles, voire confidentielles pour certaines. Il est hors de question de demander à une IA d'effectuer un quelconque traitement de ces données sans les avoir anonymisées.

Solution possible : pour gagner du temps sur l'étape d'anonymisation des données, vous pouvez demander à une IA générative de créer un outil dédié à ce travail d'anonymisation. Cet outil pourra par exemple prendre la forme d'une page HTML à exécuter localement, dans votre navigateur. Cette page devra vous offrir la possibilité de désigner le fichier à anonymiser, puis de lancer cette anonymisation, et enfin, de désanonymiser le fichier après traitement par une IA.

 Un outil possible : utiliser une IA générative comme Le Chat de Mistral AI ou Chat GPT pour générer le code de l'outil d'anonymisation /désanonymisation souhaité.

II. Comment faire ?



Conseil : Accéder à l'outil Le Chat, de Mistral AI

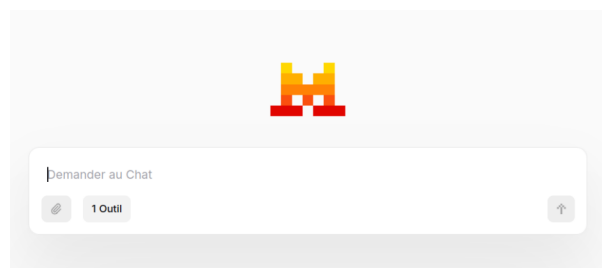
L'IA générative de Mistral AI est accessible à cette adresse : <https://chat.mistral.ai>

L'utilisation peut rester gratuite (même s'il existe des offres payantes) et peut même se faire sans création de compte.

Il est cependant conseillé de créer son compte d'utilisateur de manière à pouvoir retrouver ses différentes sessions de travail et ses précédents travaux demandés au Chat.

A. Obtenir le code HTML d'un outil d'anonymisation des données d'un fichier Excel

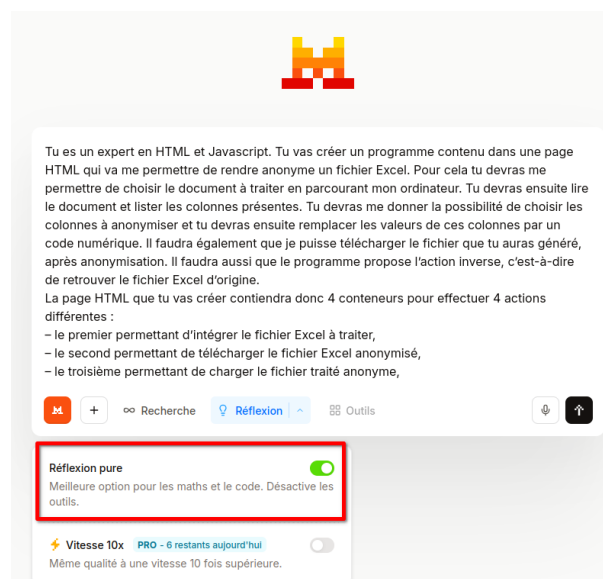
1. Accéder au Chat de Mistral AI et se connecter à son compte d'utilisateur.
2. Dans le champ de saisie, initier un nouveau chat en saisissant son prompt.



Le prompt utilisé dans cet exemple est le suivant :

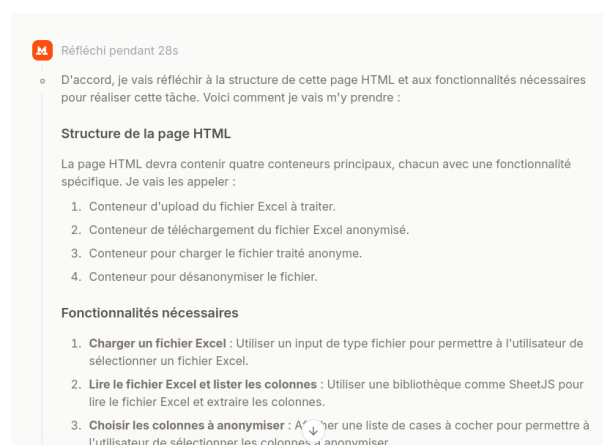
« Tu es un expert en HTML et Javascript. Tu vas créer un programme contenu dans une page HTML qui va me permettre de rendre anonyme un fichier Excel. Pour cela, tu devras me permettre de choisir le document à traiter en parcourant mon ordinateur. Tu devras ensuite lire le document et lister les colonnes présentes. Tu devras me donner la possibilité de choisir les colonnes à anonymiser et tu devras ensuite remplacer les valeurs de ces colonnes par un code alphanumérique à 6 caractères. Il faudra également que je puisse télécharger le fichier que tu auras généré, après anonymisation. Il faudra aussi que le programme propose l'action inverse, c'est-à-dire de retrouver le fichier Excel d'origine. La page HTML que tu vas créer comprendra donc 4 conteneurs : – le premier permettant d'intégrer le fichier Excel à traiter ; – le second permettant de télécharger le fichier Excel anonymisé ; – le troisième permettant de charger le fichier traité anonyme ; – le dernier permettant de faire le traitement inverse : à savoir retrouver les noms (désanonymiser). La présentation de ces 4 blocs devra être claire et aérée, avec de jolies couleurs différentes pour chaque conteneur. Tu feras attention à bien vérifier le bon fonctionnement de ton code, à te relire plusieurs fois, de sorte que je n'ai pas à revenir vers toi. »

3. Pour un meilleur résultat, choisir l'option « *Réflexion pure* » qui est censée donner de meilleurs résultats dans la création de code.



4. Attendre que l'IA donne sa réponse.

N.B. : Le temps de réflexion par l'IA est un peu plus long, mais dans ce cas de figure il vaut mieux privilégier la qualité du code fourni. Pendant ce temps de réflexion qui va durer quelques dizaines de secondes, le système indique ses intentions et la méthode qu'il va adopter pour générer le résultat escompté.



5. Copier tout le code source qui s'est affiché.



N.B. : En haut à droite du bloc de code, un bouton permet de copier la totalité du code généré par l'IA.

```
#anonymizedData {
  margin-top: 10px;
}
</style>
</head>
<body>
<div class="container">
<h1>Anonymisation de fichier Excel</h1>

<div id="section1" class="section">
<h2>Étape 1 : Charger le fichier Excel à traiter</h2>
<input type="file" id="fileInput" accept=".xlsx, .xls">
<button onclick="loadFile()">Charger le fichier</button>
<div id="columnList"></div>
</div>

<div id="section2" class="section">
<h2>Étape 2 : Télécharger le fichier Excel anonymisé</h2>
<button onclick="downloadAnonymizedFile()">Télécharger le fichier anonymisé</button>
<div id="anonymizedData"></div>
</div>

<div id="section3" class="section">
<h2>Étape 3 : Charger le fichier traité anonyme</h2>
<input type="file" id="anonFileInput" accept=".xlsx, .xls">
<button onclick="loadAnonFile()">Charger le fichier anonyme</button>
<div id="loadedAnonData"></div>
</div>

<div id="section4" class="section">
<h2>Étape 4 : Désanonymiser le fichier</h2>
<button onclick="deanonymizeFile()">Désanonymiser le fichier</button>
<div id="deanonymizedData"></div>
</div>

</div>

<script src="https://cdnjs.cloudflare.com/ajax/libs/xlsx/0.18.9/xlsx.full.min.js"></script>
<script src="https://cdnjs.cloudflare.com/ajax/libs/FileSaver.js/2.0.5/FileSaver.min.js"></script>
<script>
let workbook;
let columns = [];
let anonymizedData = [];
let originalData = [];
```

! Attention : Transposer la manipulation à d'autres types de fichiers

En modifiant le prompt, il sera possible de **transposer cet exemple** (basé sur des fichiers Excel) pour faire en sorte que l'IA génère un outil permettant de traiter d'autres types de fichiers (PDF, traitement de texte, texte brut, .csv, etc.)



Conseil : Que faire avec ce code HTML généré ?

Le travail n'est pas encore terminé, il faut maintenant arriver à utiliser l'outil qui fonctionnera grâce à ce code généré par l'IA.

Le code généré copié grâce au bouton dédié sur l'interface du Chat n'est pour le moment que du texte brut. Il va falloir maintenant « fabriquer » sur son ordinateur le fichier HTML qui contiendra l'outil souhaité.

Pour cela il faut continuer à suivre les étapes de ce pas à pas, ci-dessous.

B. Créer le fichier HTML permettant d'utiliser l'outil d'anonymisation

1. Coller tout le code précédemment copié **dans un éditeur de texte brut** comme le bloc-notes de Windows (Notepad), ou mieux dans l'incontournable éditeur de texte brut « amélioré » Notepad++ : <https://notepad-plus-plus.org/downloads>



N.B. : Si l'éditeur de code prend en charge la coloration syntaxique, le contenu du fichier pourra apparaître comme sur la capture d'écran ci-contre. Si ce n'est pas le cas, aucun problème, cette coloration sert simple-

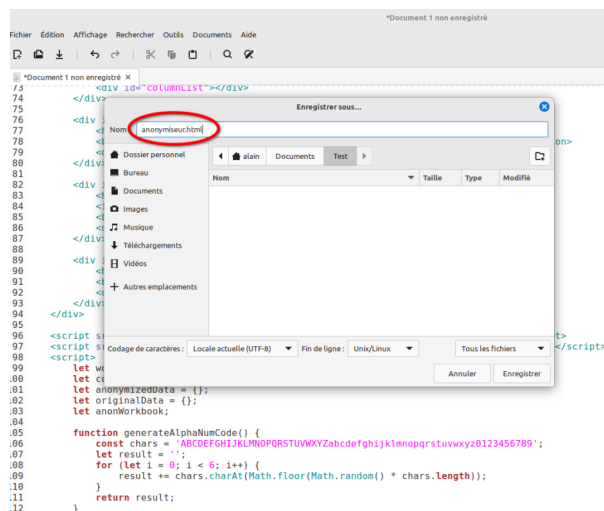
ment à avoir un meilleur confort visuel pour la lecture du code.

```

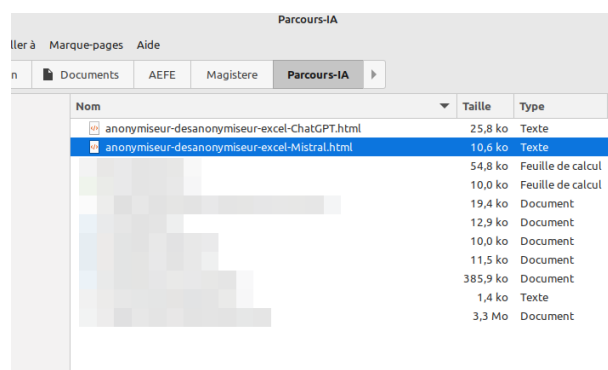
74
75
76 <div id="section2" class="section">
77   <h2>Étape 2 : Télécharger le fichier Excel anonymisé</h2>
78   <button onclick="downloadAnonymizedFile()">Télécharger le fichier anony
79   <div id="anonymizedData"></div>
80 </div>
81
82 <div id="section3" class="section">
83   <h2>Étape 3 : Charger le fichier traité anonyme</h2>
84   <input type="file" id="anonFileInput" accept=".xlsx, .xls">
85   <button onclick="loadAnonFile()">Charger le fichier anonyme</button>
86   <div id="loadedAnonData"></div>
87 </div>
88
89 <div id="section4" class="section">
90   <h2>Étape 4 : Désanonymiser le fichier</h2>
91   <button onclick="deanonymizeFile()">Désanonymiser le fichier</button>
92   <div id="deanonymizedData"></div>
93 </div>
94
95
96 <script src="https://cdnjs.cloudflare.com/ajax/libs/xlsx/0.18.5/xlsx.full.min.js">
97 <script src="https://cdnjs.cloudflare.com/ajax/libs/FileSaver.js/2.0.5/FileSaver.js">
98 <script>
99   let workbook;
100   let columns = [];
101   let anonymizedData = {};
102   let originalData = {};
103   let anonWorkbook;
104
105   function generateAlphaNumCode() {
106     const chars = 'ABCDEFGHIJKLMNOPQRSTUVWXYZabcdefghijklmnopqrstuvwxyz0123456789';
107     let result = '';
108     for (let i = 0; i < 6; i++) {
109       result += chars.charAt(Math.floor(Math.random() * chars.length));
110     }
111     return result;
112   }
113
114   function loadFile() {
115     const fileInput = document.getElementById('fileInput');
116     const file = fileInput.files[0];
  
```

2. « Enregistrer sous... » le fichier sur son ordinateur en le renommant éventuellement, mais en modifiant impérativement le format : à la place de l'extension .txt sans doute proposée par défaut, indiquer l'extension .html

→ On va donc devoir modifier dans le nom proposé pour l'enregistrement l'extension de fichier .txt par .html



3. Retrouver le fichier sur son ordinateur.

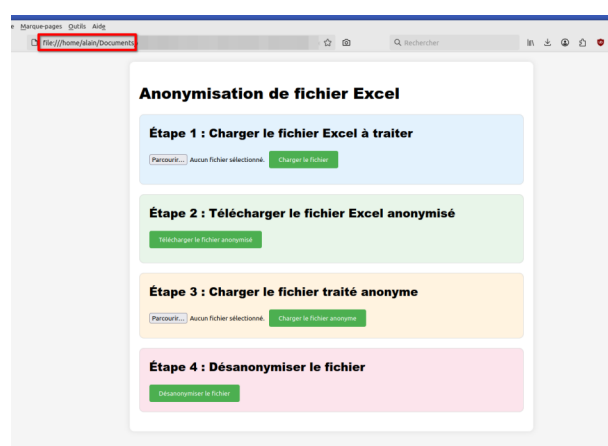


4. Double-cliquer dessus pour le lancer.

Comme pour tout fichier HTML, il s'ouvrira par défaut dans votre navigateur préféré.

N.B. : Comme pour tout fichier HTML, il s'ouvrira dans le navigateur par défaut.

À savoir : le traitement effectué se déroule entièrement localement, dans son navigateur, sur son ordinateur, aucune donnée n'est envoyée sur un serveur.



Exemple : L'outil créé par l'IA

Voilà à titre d'exemple le **fichier contenant le code HTML créé par l'IA**.

Il est possible de :

- **voir le code** en cliquant sur ce lien : [anonymiseur-desanonymiseur-excel-Mistral.txt](#)^[p.]
- **le tester** directement en cliquant sur ce lien : [anonymiseur-desanonymiseur-excel-Mistral.html](#)^[p.]

→ Pour le rendre « fonctionnel », il suffit de l'« Enregistrer sous... » en remplaçant l'extension **.txt** par **.html** puis l'ouvrir avec son navigateur.

Attention : Comment utiliser cet outil ?

Maintenant que l'outil est prêt, il ne reste plus qu'à le tester pour anonymiser/désanonymiser le contenu d'un fichier Excel.

Rien de plus simple, les étapes sont décrites ci-dessous.

C. Utiliser son outil d'anonymisation

Pour tester rapidement son outil, il est possible d'utiliser ce fichier de faux élèves (généré par une IA) : [faux_eleves.xlsx](#)^[p.]

NOM	Prénom	Classe	LV1	LV2	Régime
TFYsC1	j74hP1	3e B	allemand	italien	externe
SaEej9	jfOntM	6e A	espagnol	chinois	interne
tZ3Wnf	22nCLQ	5e B	anglais	chinois	interne
IzS6L9	ykytnq	3e D	allemand	anglais	demi-pensionnaire
z9Axc	GPXPLS	6e F	espagnol	anglais	interne
Iznhkp	j74hP1	4e A	italien	allemand	externe
mjlYqz	Bvc7zv	4e B	espagnol	italien	demi-pensionnaire
lwpSif	NPdyR	5e C	italien	anglais	externe
RhtVQk	bxLJS8	3e E	anglais	espagnol	interne
42dDpc	gyp1IS	5e C	anglais	espagnol	demi-pensionnaire
3k05Gf	rnvW4a	3e F	anglais	russe	interne
9ABUy9	j74hP1	5e F	espagnol	italien	externe
lflwnl	GPXPLS	5e B	russe	allemand	externe
LVL8CL	ilnzLt	5e A	espagnol	anglais	interne
3HsVJI	bxLJS8	3e D	chinois	russe	interne
rO79LC	zqJlxr	3e A	espagnol	allemand	interne
ANXOV4	Bvc7zv	5e F	italien	anglais	interne
XDToGh	L98tFN	4e A	chinois	russe	externe
AZuBvo	ykytnq	3e C	russe	espagnol	interne
gzIHAE	GPXPLS	4e A	anglais	russe	interne
GRa5es	E9KKP8	3e C	italien	espagnol	externe
vun7m5	257gV8	4e A	russe	anglais	interne
ISVpke	h8uekw	3e C	russe	allemand	externe
YphtdT	wyo9mh	5e B	anglais	allemand	demi-pensionnaire

Voici un aperçu de son contenu :

« Étape 1 »

Ouvrir son fichier HTML puis :

- Choisir le fichier à traiter en cliquant sur le bouton « Parcourir... » ou « Choisir un fichier » (selon votre navigateur).
- Le charger en cliquant sur le bouton vert.
- Cocher les colonnes à anonymiser.

Anonymisation de fichier Excel

Étape 1 : Charger le fichier Excel à traiter

1 Parcourir... faux_eleves.xlsx 2 Charger le fichier

Cocher les colonnes à anonymiser :

3

☒ NOM
☒ Prénom
☐ Classe
☐ LV1
☐ LV2
☐ Régime

« Étape 2 »

Cliquer sur le bouton « Télécharger le fichier [Excel] anonymisé ». Le retrouver dans son dossier habituel des téléchargements sur son ordinateur.



N.B. : Ci-contre, un aperçu de son contenu après anonymisation des noms et prénoms.

	A	B	C	D	E	F
1	NOM	Prénom	Classe	LV1	LV2	Régime
2	TFYsC1	j74hP1	3e B	allemand	italien	externe
3	SaEej9	jfOntM	6e A	espagnol	chinois	interne
4	tZ3Wnf	22nCLQ	5e B	anglais	chinois	interne
5	IzS6L9	ykytnq	3e D	allemand	anglais	demi-pensionnaire
6	z9Axc	GPXPLS	6e F	espagnol	anglais	interne
7	Iznhkp	j74hP1	4e A	italien	allemand	externe
8	mjlYqz	Bvc7zv	4e B	espagnol	italien	demi-pensionnaire
9	lwpSif	NPdyR	5e C	italien	anglais	externe
10	RhtVQk	bxLJS8	3e E	anglais	espagnol	interne
11	42dDpc	gyp1IS	5e C	anglais	espagnol	demi-pensionnaire
12	3k05Gf	rnvW4a	3e F	anglais	russe	interne
13	9ABUy9	j74hP1	5e F	espagnol	italien	externe
14	lflwnl	GPXPLS	5e B	russe	allemand	externe
15	LVL8CL	ilnzLt	5e A	espagnol	anglais	interne
16	3HsVJI	bxLJS8	3e D	chinois	russe	interne
17	rO79LC	zqJlxr	3e A	espagnol	allemand	interne
18	ANXOV4	Bvc7zv	5e F	italien	anglais	interne
19	XDToGh	L98tFN	4e A	chinois	russe	externe
20	AZuBvo	ykytnq	3e C	russe	espagnol	interne
21	gzIHAE	GPXPLS	4e A	anglais	russe	interne
22	GRa5es	E9KKP8	3e C	italien	espagnol	externe
23	vun7m5	257gV8	4e A	russe	anglais	interne
24	ISVpke	h8uekw	3e C	russe	allemand	externe
25	YphtdT	wyo9mh	5e B	anglais	allemand	demi-pensionnaire

Utiliser le fichier anonymisé pour le soumettre à l'IA pour le traitement souhaité.



N.B. : Ainsi, aucune donnée personnelle ou confidentielle n'ira alimenter l'IA utilisée.

« Étape 3 » et « Étape 4 »

Si besoin, pour effectuer le traitement inverse, c'est-à-dire pour désanonymiser les colonnes voulues, charger le fichier traité anonymisé et le désanonymiser via les boutons présents.

III. Ressources associées

Le b.a.-ba du prompt

L'intelligence artificielle est aujourd'hui quasiment omniprésente et les outils les plus plébiscités par le public sont les chatbots (comme ChatGPT ou Le Chat de Mistral AI, par exemple) qui transforment radicalement nos pratiques avec les outils numériques.

Le guide « Le b.a.-ba du prompt pour maîtriser l'IA » de Frédéric Cavazza, très synthétique, très bien fait, a pour objectif de présenter les bases et les connaissances à avoir pour s'initier à l'art de « prompter », une compétence clé pour dialoguer efficacement avec les intelligences artificielles génératives et exploiter au mieux leur potentiel.

Il aborde les points suivants :

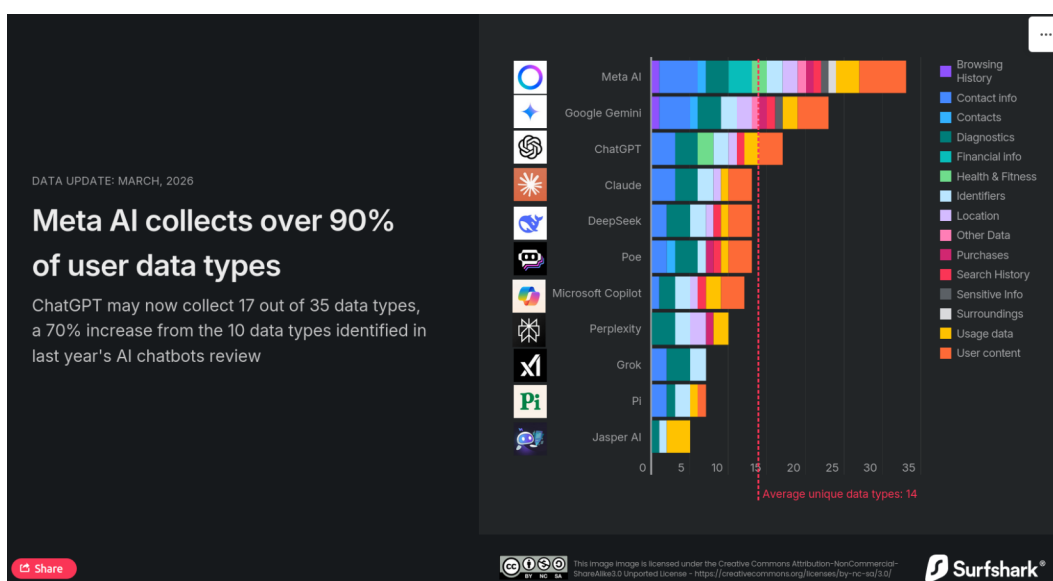
- Comprendre le fonctionnement des chatbots
- Bien appréhender la nature des prompts (leur intérêt)
- Découvrir différentes méthodes pour structurer vos prompts
- Se familiariser avec ces méthodes grâce à des exemples
- Savoir dans quelle situation utiliser quelle méthode

Ce guide est accessible depuis la page : <https://fredcavazza.net/2025/01/20/la-maitrise-des-prompts-est-une-etape-indispensable-a-ladoption-de-lia-generative/>

IV. Recommandations d'usage et juridiques

! Attention : Collecte de données à caractère personnel

On prêtera une attention toute particulière aux modèles d'IA qui seront choisis pour effectuer la tâche demandée. En effet, les différents modèles d'IA collectent tous – à des degrés divers – des données des utilisateurs et certains sont beaucoup plus gourmands que d'autres ! Voir le comparatif (récent) sur l'image ci-dessous où on voit bien la différence entre les modèles qui collectent le moins de données (au bas du graphique) et ceux qui en collectent le plus (en haut du graphique).



Surfshark, « Meta AI, Google Gemini, and ChatGPT are top data collectors »

Par exemple, on évitera absolument d'utiliser l'IA de Meta ou encore Gemini de Google. Faut-il s'en étonner, de la part d'outils proposés par ces entreprises ?

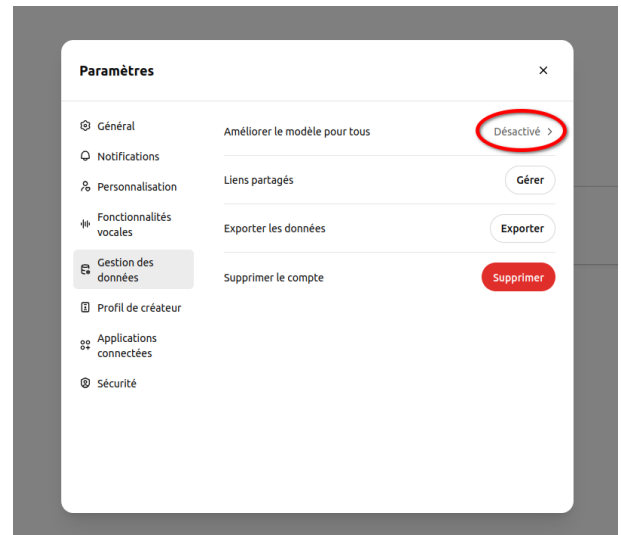
📎 Conseil : Protection des données fournies

Certains modèles d'IA génératives s'enrichissent grâce aux données fournies par les utilisateurs ainsi qu'aux données des conversations avec les chatbots.

Il est essentiel de garder cela à l'esprit quand on décide de demander à une IA de travailler sur la base de nos propres documents : on veillera absolument à « préparer » les contenus qui seront fournis à l'IA pour les anonymiser au maximum, pour les débarrasser de toute référence personnelle, AVANT de les faire traiter par l'IA.

Il est également possible dans certains cas de signaler à l'outil choisi qu'on ne souhaite pas que nos données puissent servir à l'entraîner.

Par exemple, dans ChatGPT, on peut aller dans Paramètres > Gestion des données, et désactiver « Améliorer le modèle pour tous ».



Ou bien, dans Le Chat de Mistral AI, il faut aller dans la partie administration de son compte utilisateur, puis sur cette page des préférences : <https://admin.mistral.ai/chat/preferences> et décocher la case « Autoriser vos interactions à être utilisées pour entraîner nos modèles ».

Même si l'outil précise que « Tout ce qui est partagé est anonymisé », il reste préférable de ne pas laisser le modèle d'IA s'enrichir à partir de nos données.

