

Nom :	Prénom :	Classe :
-------	----------	----------

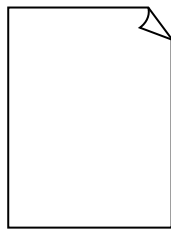
SNT — Métadonnée, données personnelles

Élément du programme : Les données structurées et leur traitement → Métadonnées. Retrouver les métadonnées d'un fichier personnel. Données → Définir une donnée personnelle.

Métadonnée ?

Les **métadonnées** sont des **données** qui décrivent les propriétés, le contenu ou la structure d'un ou plusieurs fichiers. Elles s'ajoutent aux **données brutes** en fournissant des informations complémentaires.

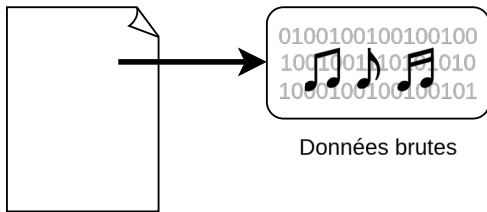
Prenons par exemple ce fichier `morceau.mp3` :



morceau.mp3

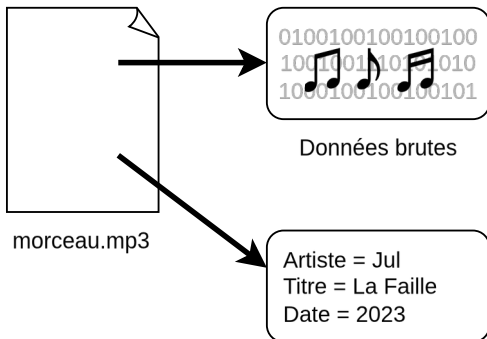
Le **nom de ce fichier** est `morceau`,
et son **extension** est `mp3`.

Cette extension appartient à la catégorie des fichiers **audio**¹.



morceau.mp3

La chanson enregistrée (encodée) dans ce fichier forme ce que l'on appelle les **données brutes**.



morceau.mp3

Métadonnées

Les informations liées à cette chanson forment ce que l'on appelle les **métadonnées**.

Dans cet exemple, les métadonnées nous permettent de voir qu'il s'agit d'un morceau de 2023, intitulé « La Faille », et chanté par l'artiste « Jul ».

¹ Relire à ce propos le cours « Formats de données, données structurées, données sur le cloud ».

► Exercice 1 (à vous de jouer)

Dans le tableau ci-dessous, les métadonnées données en exemple sont les suivantes :
Date, géolocalisation, auteur et appareil.

Cochez les cases des métadonnées qui sont, selon vous, présentes dans les fichiers proposés :

Nom du fichier	Date	Géolocalisation	Auteur	Appareil
Umbrella.mp3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Astérix.mp4	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
IMG_2912.jpg	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
cv.docx	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Données personnelles

D'après la Commission Nationale de l'Informatique et des Libertés (CNIL) — un organisme indépendant garant du **respect de la vie privée** et de la **protection des données personnelles** en France — on appelle **donnée personnelle** « toute information identifiant directement ou indirectement une personne physique (ex. nom, numéro d'immatriculation, n° de téléphone, empreinte digitale, etc.) ».

► Exercice 2

Rayez les mots qui ne relèvent pas de la définition de la *donnée personnelle* de la CNIL :

Prénom	Iris	Tarif	Âge
Siège social	Météo	Empreinte digitale	Logo

Risques liés à l'utilisation et à la divulgation des métadonnées

Les risques liés à l'utilisation et à la divulgation des métadonnées sont nombreux et importants, notamment en matière de confidentialité et de sécurité. Voici une liste des principaux risques :

- Divulgence d'informations sensibles : Les métadonnées peuvent contenir des informations cachées, comme l'auteur d'un document, la date et le lieu de création, ou encore les coordonnées GPS d'une photo. Leur divulgation accidentelle peut révéler des données confidentielles ou personnelles non prévues à destination des tiers.
- Exposition de réseaux de communication : L'analyse des métadonnées d'échanges (emails, messages instantanés) peut révéler qui communique avec qui et quand, exposant ainsi des relations professionnelles, juridiques ou privées.
- Exploitation commerciale ou politique : Hors contrôle, les métadonnées peuvent être exploitées pour le profilage, la publicité ciblée, ou la création de **bulles de filtre**². Les métadonnées peuvent contenir des informations indirectes sur des opinions, lieux fréquentés, habitudes, rendant possible un suivi invasif ou des discriminations.

2 Les bulles de filtre (également appelées bulles cognitives) désignent un phénomène où une personne est **enfermée** dans une **sphère d'information** principalement constituée de contenus **qui confirment ses propres opinions**, croyances ou comportements, grâce à des algorithmes et biais cognitifs.

► Exercice 3

Prenez connaissance du document ci-dessous, puis répondez aux différentes questions posées.

Piratage de France Travail : 43 millions de victimes

C'est la plus importante cyberattaque de l'histoire de France Travail et même l'une des plus graves à l'échelle française, en termes de quantité de données compromises. Le piratage ayant visé l'ex-Pôle Emploi en début d'année a permis aux hackers d'accéder à 43 millions d'identités. Celles-ci comprenaient : nom, prénom, adresse, courriel, numéro de téléphone et numéro de sécurité sociale. Un véritable bottin à ciel ouvert comprenant plus de la moitié des Français.

[...] Notre enquête révèle par ailleurs que France Travail n'en est pas à sa première cyberattaque. Au moins deux autres piratages ont visé l'organisme ces dernières années, posant là aussi la question de sa responsabilité. En 2021, une première attaque a été causée par la négligence d'un salarié qui avait pour habitude de télécharger de grandes quantités de données pour les mettre à disposition d'organismes extérieurs. L'un de ces fichiers, qui s'est retrouvé sur le serveur non sécurisé d'une agence d'intérim, a été volé. Des articles de presse se font alors l'écho d'une fuite d'un peu plus d'1,5 million de données. France Travail reconnaît l'intrusion mais ramène ce chiffre à environ 50 000 identités.

[...] Ce type de piratage présente des risques, et notamment celui du phishing, qu'on appelle "hameçonnage" en français. Il s'agit de ces messages frauduleux que l'on reçoit désormais fréquemment par SMS ou par mail, dans lesquels un escroc tente de se faire passer pour ce qu'il n'est pas : une banque, un service de livraison... Ou ici, France Travail. "*L'attaquant dispose d'informations exactes (nom, adresse, etc.) qui mettent en confiance la victime. Celle-ci peut penser que le message qu'elle reçoit est authentique*", explique Thibaud Antignac, adjoint au chef de l'expertise technologique à la CNIL, l'organisme chargé de protéger nos données personnelles.

Extraits de l'article France Info rédigé en 2024 par Maxime Tellier, avec la cellule investigation de radiofrance.

Questions	Vos réponses (à compléter)
Comment définiriez-vous le terme <i>phishing</i> ?	
Quelles sont les noms des six données personnelles dérobées dans ce piratage de France Travail ?	
Comment une personne malintentionnée pourrait-elle utiliser les données dérobées ?	
Selon vous, est-ce que l'IA pourrait avoir un impact dans ce type d'affaire ?	

Pour aller plus loin

Pour poursuivre la découverte des métadonnées, vous pouvez essayer d'envoyer une photo préalablement prise par votre téléphone à l'adresse suivante :

<https://fmr.tf/metadata/>



Cette page affiche toutes les métadonnées présentes dans une image.