MANUEL SUR L'ARCHIVAGE EN LIGNE ET CONSERVATION DES PREUVES EN LIGNES.

Internet n'est pas une archive ! La durée de vie moyenne d'une page web est de 92 jours.

Il existe des pratiques d'archivages web qui permettent de conserver l'historique d'une page et de ses métadonnées. Ce processus n'est pas automatisé et il faut **manuellement faire archiver ces pages.**

À travers l'outil Wayback, un utilisateur peut

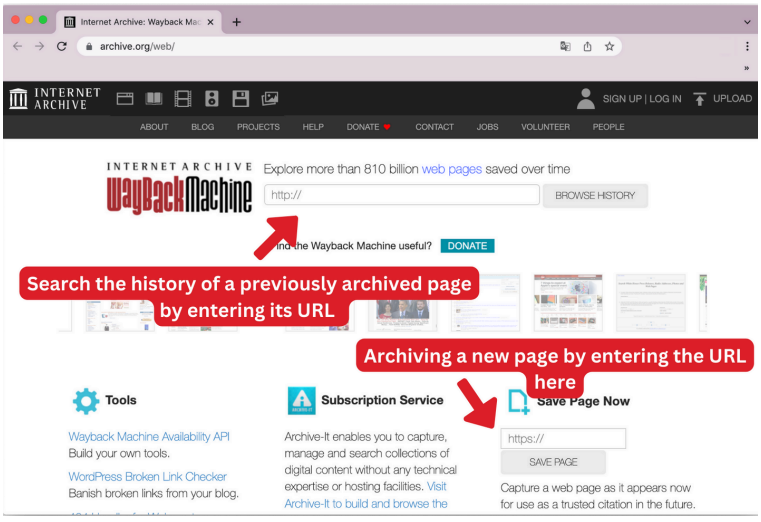
- faire archiver une page en un clic
- remonter l'historique des pages archivées et accéder aux anciennes versions de cette page —si elles ont été archivé !

Même si l'URL dirige l'utilisateur vers une page n'existant plus (voir annexe 1), Wayback est à même de récupérer les informations de cette page si elle a été archivée.

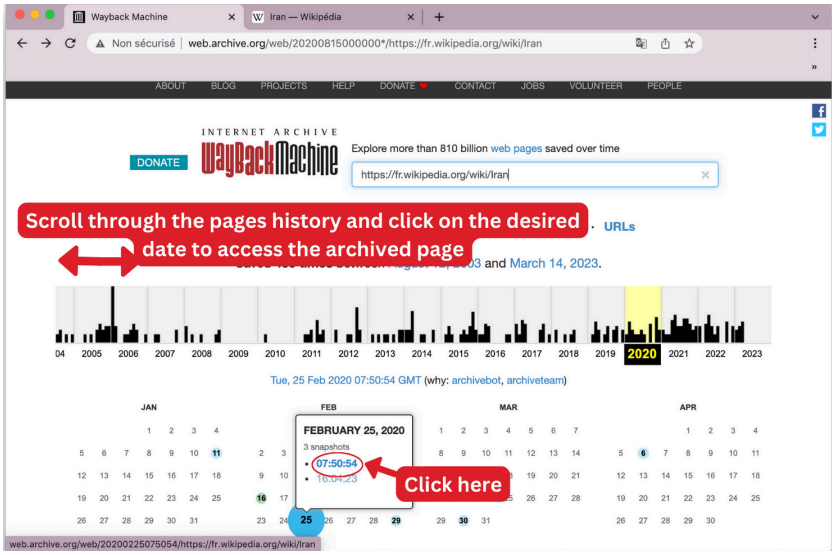
Archiver : Utiliser Wayback Machine : http://web.archive.org/

1) Vérifier que l'URL n'a pas déjà été archivée à travers la barre de recherche. Si l'URL n'existe pas entrer la nouvelle URL sous la rubrique "Save Page Now".

Attention : Pour archiver une vidéos ou une photo sur twitter, veiller à ce que les paramètres du comptes permettent la visualisation d'images sensibles.

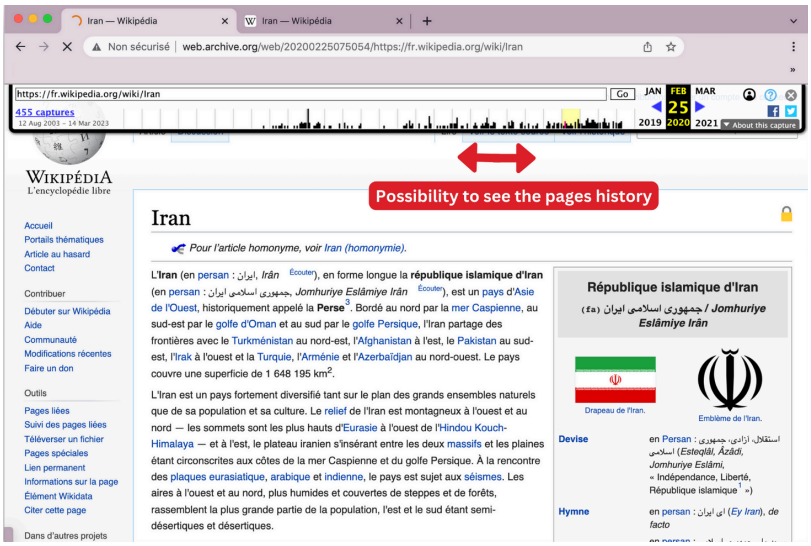


Il existe egalement d'autres sites d'archivage très performant comme : https://archive.ph/



2) Une fois l'URL enregistrée, l'historique de la page s'affichera et permettra de visualiser les modifications apportées à la page avec des dates et heures précises de modifications.

3) En cliquant sur une date précise, on peut visualiser la page telle qu'elle a été capturée. On peut également utiliser la barre chronologique afin de faire défiler une page web et observer son évolution.



Stocker sur vos appareils

Stocker de manière sécurisé

Il est préférable de ne pas stocker de preuve importante sur des appareils reliés au cloud. Si les informations ont besoin d'être partagées à travers internet, il est préférable de **passer par des clouds et des adresses email cryptés comme proton** [voir : https://proton.me/]. Proton contrairement à Google et Outlook crypte les informations avant même de les enregistrer sur le cloud, ce qui signifie que le serveur n'a pas accès aux informations stockées.

Attention : Dans le cas où vous utilisez un cloud, il faut s'efforcer de toujours sauvegarder les informations sur un ordinateur et de f**aire des sauvegardes régulières** sur un disque dur afin de ne pas perdre les informations. Il est également plus prudent d'**utiliser un email, mot de passe et compte dédié au stockage d'information,** et pas une adresse personelle ou usuelle.

Il est parfois impossible de télécharger les vidéos, dans ce genre de situation on peut avoir recours à Excel afin de consigner les URL des preuves et les nommer (ne pas oublier de faire archiver ces pages !). Pour rappel, Microsoft Excel ainsi que Google Sheet qui sont les deux plateformes en ligne les plus utilisées pour la création de fichiers Excel ne sont pas des plateformes cryptées. Il faut donc veiller à **limiter les options de partages** et favoriser **l'utilisation de l'application local de excel sur votre appareil** et si besoin sauvegarder une copie sur un drive crypté.

> **Attention :** Lorsque l'on est amené à travailler sur des **réseaux wifi publics**, l'activité peut être facilement espionnée. En particulier les wifis publics de gare et d'aéroport. **Munissez-vous d'un VPN** afin de protéger votre activité en ligne et vos mots de passe !

Sur le téléphone

Il est possible de télécharger des vidéos et des informations à partir d'un téléphone. Tous les logiciels sous Apple iOS et Google Android n'offrent pas la possibilité de nommer une vidéo dans la galerie du téléphone. Il est donc souhaitable de stocker les preuves et vidéos sur une application drive [proton est accessible sur Apple et Android] afin d'y mettre les informations et les nommer correctement.

Conserver les preuves

Plus une vidéo est manipulée plus ses métadonnées sont complexes. S'il faut partager la vidéo et l'analyser (la faire passer d'appareil en appareil) ou la changer de format (passer de mp4 à MOV), toutes ces actions peuvent porter préjudices à la recevabilité de la preuve dans le cadre légal. **Le protocole de Berkeley préconise de copier la vidéo afin de l'analyser et de conserver la version original intacte.** Cela est valables pour toute forme de fichier (vidéo, image, format texte, etc.)

Nommer

Archiver une preuve ce n'est pas seulement la conserver c'est aussi faire en sorte qu'elle puisse être facilement citée, il est donc vital d'étiqueter les preuves :

yyyy-mm-dd_CreatorName_IncidentDescription.
année-mois-date_NomDeLaPersonneQuiFilm_DescriptionDeLincident

Lors d'une collecte, il faut systématiquement faire :

- Une capture d'écran de la preuve avec la date et l'heure rendue visible ainsi que les informations de publication (nom de l'utilisateur)
- Conserver l'URL de la page où se situe la page web et l'archiver.
- Si du contenu multimédia est inclus dans la preuve télécharger ce contenu séparément en nommant et stockant correctement et de manière sécurisée.

Citer une page web et des fichiers en ligne

Dans la recherche

Internet n'est pas une archive ! Pour s'appuyer sur une page web et les informations qu'elle contient dans le cadre de la recherche les informations et la page web doivent rester identiques afin de ne pas changer le matériel de la recherche.

Dans le droit

D'après la Cour Pénal Internationale, l'admissibilité d'une preuve audiovisuelle tirée d'internet (une preuve en open source) repose sur trois facteurs. La pertinence de la preuve, la valeur probante de la preuve et l'absence d'effet préjudiciable. Tel qu'il est expliqué dans les statuts de Rome sur le traitement des preuves, **une preuve tirée de l'open source doit être soumise à une évaluation des métadonnées** pour que la véracité soit établie.

Tout dossier reposant sur des preuves collecté en open source tel que des preuves audiovisuelles encourt des risques si les informations en ligne venaient à être modifiées, abimées ou corrompues.

> **Attention :** Les vidéos floutées ou comportant des sous-titres sont considérées comme corrompues. Il faut s'efforcer de trouver la source de la vidéo originale.

Formulaire de collecte d'information en Open Source

Afin que les métadonnées d'une preuve en Open Source puissent être collectés, les conditions de collecte de la preuves sont importantes. Le formulaire proposé par le **Protocole de Berkeley** (voir annexe 2) stipule que la date et heure de la collecte de l'information, ainsi que l'identité de la personne collectant la preuve et l'adresse IP de l'appareil sur laquelle la preuve a été collecté doivent être systématiquement stipulées.

Voir le formulaire proposé par le protocole de Berkeley

Annexe p.86

ANNEX IV

Annex IV

Online data collection form	
1. Collector information	
Investigation:	
Collector:	
Collector IP address:	
Start of collection (date/time stamp):	
End of collection (date/time stamp):	
2. Target information	
Web address (URL):	
HTML source code:	
Screenshot:	
Captured data:	
IP address(es):	
3. Collection package information	
Collection package file name:	
Collection package hash list:	
Hash of collection package hash list file:	
4. Services used	
Software product(s):	
Time service:	
IP service:	
WHOIS service:	

Annexe 1 :

Pages Web

Construction d'une page web

Une page web est un support à travers sont rendus accessibles des fichiers ou informations. Une page web est accessible par une URL et comporte des **métadonnées**.

URL ou "Uniform Ressource Locator" est une adresse web qui permet de se diriger vers une information sur internet. Une URL ne garantit pas la qualité de l'information et est seulement un système de notation permettant d'y accéder.

Métadonnées : Les métadonnées sont des informations rattachées à des informations/fichier en ligne. Si une lettre est une donnée, les métadonnées de la lettre sont la date d'envoi, le timbre, le poids de la lettre et l'adresse de l'expéditeur. Sur internet toutes ces informations peuvent également être tracées.

Une page Web dans le temps

Une page web peut se dégrader dans le temps au point de ne plus contenir les mêmes informations. La durée de vie moyenne d'une page web est de 4 ans. Ainsi une URL peut mener les utilisateurs a une page ne comportant plus les mêmes informations.

Les informations peuvent avoir été déplacées, l'architecture du serveur peut avoir été modifiée, ou encore le dossier et les informations en ligne peuvent avoir été délibérément attaqués et abimés. Dans certains cas des virus peuvent s'attaquer de manière ciblée à des informations.

Annexe 2 :

Protocole Berkeley

Avec l'usage grandissant des réseaux sociaux comme outil informatif et militant, internet est devenu un lieu éminemment politique qui permet la propagation d'information, de photos et de vidéos. Ces informations peuvent être utiles dans un cadre légal et utilisées comme preuve, c'est à cet égard que les Nations Unies ont publié en 2022 le *Protocole de Berkeley sur la recherche digitale en Open Source.*

Ce protocole cherche à uniformiser les pratiques de collecte d'information en ligne afin qu'elles soient exploitables par l'IPC ainsi que d'autres instances internationales. Ce manuel est conforme aux recommandations du protocole.

Bibliographie :

Alexa Koenig et Mehandru Nikita. « Open Source Evidence and the International Criminal Court », 2019. https://harvardhrj.com/2019/04/open-source-evidence-and-the-international-criminal-court/.

Neal, James G. « The Integrity of Research Is at Risk: Capturing and Preserving Web Sites and Web Documents and the Implications for Resource Sharing ». IFLA, 3 juillet 2014, 24.

Nielsen, Janne. « Using Web Archives in Research – an Introduction ». NetLab, 2015.

Peyssard, Jean-Christophe. « Archivage des données du web : sauvegarde et citabilité pour la recherche ». CNRS, juillet 2022. https://distam.hypotheses.org/3931

Teszelszky, Kees. « Introduction: Digital Humanities and the Use of Web Archives ». International Journal of Digital Humanities 2, no 1-3 (novembre 2021): 1-4. https://doi.org/10.1007/s42803-021-00040-5.