

Mener une recherche d'informations

Actualisé le 03/11/2023

Université Paris-Saclay : Les Bibliothèques
2023-2024

Objectifs de la séance :

- Pouvoir trouver la bonne information sur Internet et les ressources académiques



Quand on choisit un livre à la bibliothèque, on ne se pose généralement pas la question de la validation de l'information car l'ouvrage a été sélectionné par un bibliothécaire, après qu'il l'ait été par un éditeur...

Il en est de même pour les articles publiés dans les revues scientifiques. Ils sont la plupart du temps soumis à un comité de lecture (on parle alors de « relecture par les pairs », c'est-à-dire une relecture de l'article par un comité de chercheurs qui valident la démarche et le contenu scientifique du document.)

La notion de validité est différente sur internet car la plupart du temps, la validation est inexistante.

Sur Internet on trouve tout et n'importe quoi !

Cas particulier de Wikipédia : c'est une encyclopédie collective, chacun peut corriger et modifier mais les auteurs ou modificateurs ne sont pas toujours des experts !

Les questions à se poser

- Où ? Comment est construite l'URL...

Elle est constituée de la manière suivante :

Protocole://Nom de serveur.Nom de Domaine
/Répertoire/Document

En pratique...

<https://vaccination-info-service.fr/Les-maladies-et-leurs-vaccins/Infections-a-Papillomavirus-humains-HPV>

Université Paris-Saclay : Les Bibliothèques 2023-2024

université
PARIS-SACLAY

A partir des résultats d'une recherche, le premier élément à regarder est **l'URL : c'est l'adresse du site**, qui fournit des informations primordiales pour pouvoir évaluer l'information qu'on y trouvera, connaître l'auteur ou l'organisme auteur du site, etc.

L'URL consiste en : Protocole://Ordinateur/Répertoire/Document.NomDeDomaine

Protocole = http, ftp...

- **Ordinateur** = Hébergeur du site, y être très attentif, c'est souvent là que l'on peut voir s'il s'agit du nom d'une entreprise, d'un organisme public, d'un forum, d'une page perso...
- **Répertoire/document** = structuration interne du site (=dossier, sous-dossiers, et document, comme sur votre ordinateur)
- **Nom de domaine** = .net, .fr, .gouv etc. donne une indication, mais les noms de domaines peuvent être achetés. (1) (2)

(1) Cas pratique sur les arnaques aux cartes grises : <https://www.eplaque.fr/carte-grise/carte-grise-gouv.html>

(2) Pour aller plus loin sur l'évaluation d'une URL : <http://pointdoc.ac-creteil.fr/IMG/old/peda/rechdoc/lycee/carnet%20de%20bord/pdf/methodo/url.PDF>

Les bonnes adresses...



<https://solidarites-sante.gouv.fr/ministere/acteurs/agences-et-operateurs/>

Université Paris-Saclay : Les Bibliothèques 2023-2024

université
PARIS-SACLAY

Les sites en « .gouv.fr » sont des sites du gouvernement français : vous y trouverez une information fiable et officielle en matière de santé.

Point sur le protocole HTTPS

Le protocole **https** est un protocole sécurisé qui garantit un cryptage des données (en particulier personnelles) que vous transmettez au site visité. Cela évite notamment qu'un pirate intercepte ces données entre vous et le site (« Man-in-the-middle attack » (1)). Cela est indispensable notamment dans le cas de données sensibles telles que des données bancaires, relatives à votre identité ou encore à votre santé, ou tout simplement pour vos identifiants et mots de passe.

Ainsi, il est fréquent que des pirates proposent des copies presque parfaites de sites institutionnels ou commerciaux (banques, assurances, sécurité sociale, etc.). Un moyen de les reconnaître est de scruter attentivement leur URL. L'usage d'un protocole http simple (et non https) peut être un indice que le site est suspect. Soyez vigilants aussi à l'orthographe exacte de l'URL du site authentique : avec une seule lettre ou un seul signe différents, on peut vous tromper.(2) (3)

(1) https://fr.wikipedia.org/wiki/Attaque_de_l%27homme_du_milieu

(2) Mozilla et Chrome pénalisent même les sites qui ne sont pas en HTTPS (voir : <https://www.clubic.com/navigateur-internet/mozilla-firefox/actualite-873498-firefox-70-certificats-visibles-chiffrement-https.html> et <https://www.clubic.com/navigateur-internet/google-chrome/actualite-871956-chrome-79-bloquer-contenus-securises-pages-https.html>)

(3) Pour en savoir plus sur le protocole HTTPS : <https://junto.fr/blog/https/>



Sur cette diapositive, une liste non exhaustive d'organismes fiables.

Pour une bonne sélection de sites en santé, vous pouvez consulter cette page du Ministère :

<https://solidarites-sante.gouv.fr/ministere/acteurs/agences-et-operateurs/>

Lorsqu'on a sélectionné un site selon les premiers éléments (titre, URL, résumé du site) on se rend sur le site et le premier critère pris en compte dans le choix d'un site est indéniablement son apparence : organisation des pages, orthographe, graphisme, présence ou non de photos, animations trop prononcées...

L'ergonomie du site a une importance à priori et incite parfois même à ne pas aller plus loin (lisibilité des informations, accès).

POUR ALLER PLUS LOIN :

Si vous maîtrisez l'anglais, le moteur de recherche TRIP vous renverra vers des articles et des recommandations : <https://www.tripdatabase.com/>

ATTENTION, dans tripdatabase, les recommandations concernent surtout la sphère anglophone, même si on y trouve des recommandations de la HAS. Ce moteur établit un niveau de fiabilité des documents et des travaux en valorisant ceux fondés sur une « pratique fondée sur les preuves » (« Evidence based practice ») : une pyramide avec un code de couleur en face de chaque référence indique la nature du document ainsi que son niveau de fiabilité (par exemple, une recherche originale et préliminaire = « primary research » est présentée comme moins fiable qu'une revue de la littérature ou un protocole basé sur les preuves).

Les questions à se poser

- Quoi ? Que propose le site...
- Qui ?
- « Qui sommes-Nous ? »

Mentions légales

Faire une recherche.
Une aide précieuse : Decodex.

Université Paris-Saclay : Les Bibliothèques 2023-2024 UNIVERSITÉ
PARIS-SACLAY

Voici quelques critères d'évaluation importants :

- **crédibilité de la source** (l'organisation est-elle sérieuse, fiable?)
- **crédibilité de l'auteur** (quelles sont sa notoriété, son expertise sur le sujet, ses intentions?)
- **crédibilité du document** (quelle est la fraîcheur de l'information? Le document est-il clairement structuré? La rédaction et l'orthographe sont-elles de bonne qualité?)
- **crédibilité du contenu** (les sources sont-elles citées? Il y a-t-il une bibliographie?)
- **pertinence de l'information** (l'information correspond-elle vraiment au sujet de recherche, aux besoins d'informations et au niveau de connaissances requis?)

QUOI ?

Quel public visé ? Grand public ? Universitaire ?

Niveau de précisions des informations

Analyse des sources, les liens vers d'autres sites, **les sites qui ont un lien vers ce site**

(« backlinks »)(1)

QUI ?

Évaluer la notoriété de l'auteur : est-il un spécialiste du domaine ? Quelle est son activité professionnelle, son expérience ? Appartient-il à une société savante, une institution, un organisme reconnu ? Quelle est sa reconnaissance par ses pairs ? Publie-t-il sur le sujet (faire une recherche par auteur dans Google scholar) ?

Publie-t-il cet article pour son compte personnel ou au nom de son organisme de tutelle ?

Le « qui » peut être dans le bandeau d'accueil de la page, dans le « Qui sommes-nous? » ou dans les mentions légales.

Chercher donc la rubrique « **Qui sommes-nous ?** » ou « **mentions légales** ». **Si l'auteur n'est pas clairement identifié, se méfier du site.**

Les bases de type **WHOIS** sont un outil précieux : elles donnent des indices sur qui se cache derrière une URL ou une IP : <https://whois.domaintools.com/> ou pour les domaines français : <https://www.afnic.fr/fr/produits-et-services/services/whois/>

Exercice : allez sur le site « vaccination info service » et cherchez qui est responsable du contenu.

(1) Cette information en dit long sur le niveau de sérieux ou la tendance idéologique d'un site (se rappeler que la science est objective et ne doit en aucun cas être biaisée par une idéologie). Ce site est intéressant pour donner un aperçu des liens qui renvoient vers une URL qui vous intéresse : <https://www.openlinkprofiler.org>

Les questions à se poser

- Quand ?
 - Date de mise à jour
 - Liens valides
 - Bases whois

Université Paris-Saclay : Les Bibliothèques 2023-2024

université
PARIS-SACLAY

QUAND ? (1)

- Date de création de la ressource.

Celle-ci peut être visible dans les propriétés d'un document (exemple : PDF).

En ce qui concerne une page internet, la date de création peut être (rarement sauf pour les billets de blog par exemple) mentionnée dans la zone de titre, mais parfois aussi en très petit en bas ou en haut de la page. La plupart du temps cette mention est inexistante sur beaucoup de pages web.

- Fréquence et date de mise à jour (de quand date la dernière actualité ?). On rencontre ici les mêmes problèmes que dans le paragraphe précédent : parfois la date de mise à jour apparaît discrètement en bas ou en haut de la page. Cependant, le navigateur Mozilla offre la possibilité en faisant un clic droit de sélectionner « Informations sur la page » et de voir la date de mise à jour (on a parfois de drôles de surprises!).

Ne nous réjouissons cependant pas trop vite, car certaines mises à jours périodiques sont parfois le fait d'un robot qui assure la maintenance du site.

- Les liens vers d'autres sites sont-ils à jour? Existe-t-il beaucoup de liens aveugles? Maintenant, cela est moins le cas car des logiciels existent pour détecter ce type de liens (w3C link checker : <https://validator.w3.org/> , Broken Link Check : <https://www.brokenlinkcheck.com> ...).

(1) Le paradoxe du QUAND

En tous domaines et en particulier en sciences et dans le domaine de la santé en général, les informations les plus à jours, les plus récentes, donc, sont les plus valables.

Cependant, sur certains sujets d'actualité où la recherche manque de recul, la fraîcheur peut ne pas être un signe de fiabilité (manque de recul, risque de polémique), voir la mention faite dans le chapeau de certains articles de Wikipédia :

« **Cet article concerne un événement en cours.** Ces informations peuvent manquer de recul, ne pas prendre en compte des développements récents ou changer à mesure que l'événement progresse. Le titre lui-même peut être provisoire. N'hésitez pas à l'améliorer en veillant à **citer vos sources.** »

Ceci concernait par exemple la page suivante au cours des la pandémie de Covid-19 :

https://fr.wikipedia.org/wiki/Pand%C3%A9mie_de_Covid-19

Et après...

- Attention à l'utilisation des ressources = respecter le droit d'auteur
- Code de la propriété intellectuelle > Droit d'auteur
- Droit moral VS Droit patrimonial

Université Paris-Saclay : Les Bibliothèques 2023-2024

université
PARIS-SACLAY

La protection des œuvres

Même si une ressource est mise à disposition gratuitement sur le web, cela ne veut dire pour autant que l'on peut en faire ce que l'on veut. Le web n'est pas une zone de non-droits, il faut respecter la législation en vigueur : la code de la propriété intellectuelle régit le droit d'auteur. Cela concerne les œuvres matérialisées : sont exclus les idées, les concepts, les méthodes. L'auteur jouit d'un droit de propriété exclusif dès la création, sans nécessité d'accomplissement de formalités (dépôt ou enregistrement). > Toutefois, en cas de litige, il est nécessaire de pouvoir apporter une preuve de l'existence de l'œuvre à une date donnée.

Droit moral et droit patrimonial :

Droit moral : il concerne la paternité de l'œuvre et protège l'intégrité de l'œuvre : sans limite de durée : l'auteur de Roméo Et Juliette sera toujours Shakespeare!

Droit patrimonial : il permet à l'auteur d'être rémunéré pour une durée variable (en France 70 ans droits d'auteur après la mort de l'auteur, les bénéficiaires sont les héritiers ou « ayants-droits », 50 ans pour les droits voisins (1)).

Les brevets sont protégés pour 20 ans.

A l'issue de ces périodes le document tombe dans le domaine public.

Dans les pays anglo-saxons, c'est le copyright qui protège les œuvres de l'esprit.

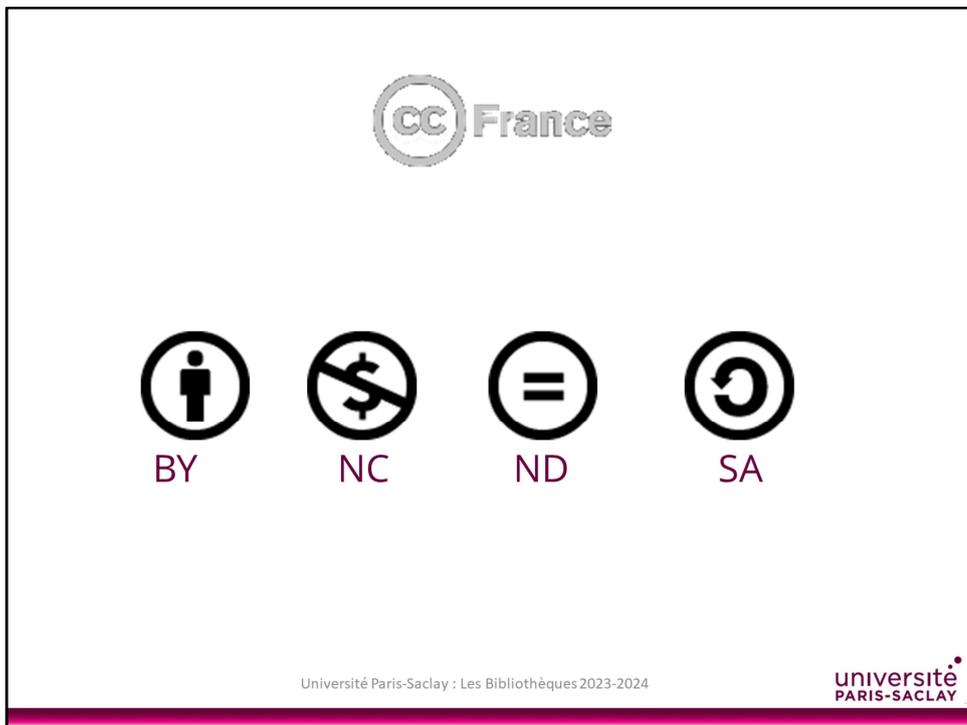
Contrairement à la France, un dépôt est nécessaire afin de le faire valoir aux Etats-Unis.

Les œuvres ayant fait l'objet d'un dépôt de copyright peuvent ainsi afficher le symbole ©, suivi de l'année de publication, puis du nom de l'auteur (ou de la société ayant déposé le copyright).

En France, les mentions copyright, ©, et « Tous droits réservés » n'ont aucune valeur juridique, et n'exclut pas de contacter l'auteur !

Attention : L'absence de sigle ou de mention du droit d'auteur ne signifie pas que l'œuvre n'est pas protégée ! Ainsi tous les éléments présents sur Internet (images, vidéos, extraits sonores, textes) sont soumis de facto au droit d'auteur, même si leur accès est libre et gratuit et qu'aucune mention ne précise qu'ils sont protégés !

(1) « Si le droit d'auteur est un droit de propriété intellectuelle sur une œuvre de l'esprit, les « droits voisins du droit d'auteur » sont accordés à des personnes physiques ou morales qui ont participé à la création de cette œuvre, mais n'en sont pas les auteurs premiers. » Le Monde avec AFP, 23 octobre 2019, https://www.lemonde.fr/pixels/article/2019/10/23/qu-est-ce-que-le-droit-voisin-au-centre-d-un-conflit-entre-les-medias-et-google_6016644_4408996.html



L'auteur peut permettre de reproduire, modifier et diffuser librement son œuvre sous réserve des conditions indiquées dans son contrat de licence.

Une des premières initiatives concerne les licences de type GNU pour les logiciels libres.

Dans la lignée, on trouve aujourd'hui le système des **creative commons**. L'initiative a pour but de diffuser et partager plus librement les créations intellectuelles, scientifiques et artistiques, notamment par Internet.

L'on n'est pas dispensé toutefois de la nécessité de toujours mentionner auteur et source, surtout dans un contexte universitaire et scientifique !

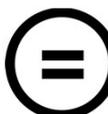
Les licences de type **creative commons** sont composables avec 4 conditions dont les différentes combinaisons donnent 6 types de licences :



« **BY** » : L'œuvre peut être librement utilisée, à la condition de l'attribuer à l'auteur en citant son nom. Cela ne signifie pas que l'auteur est en accord avec l'utilisation qui est faite de ses œuvres. (Source : Wikipedia)



Réutilisation et partage de l'œuvre autorisés mais exactement sous les mêmes conditions que celles choisies à l'origine par son auteur.

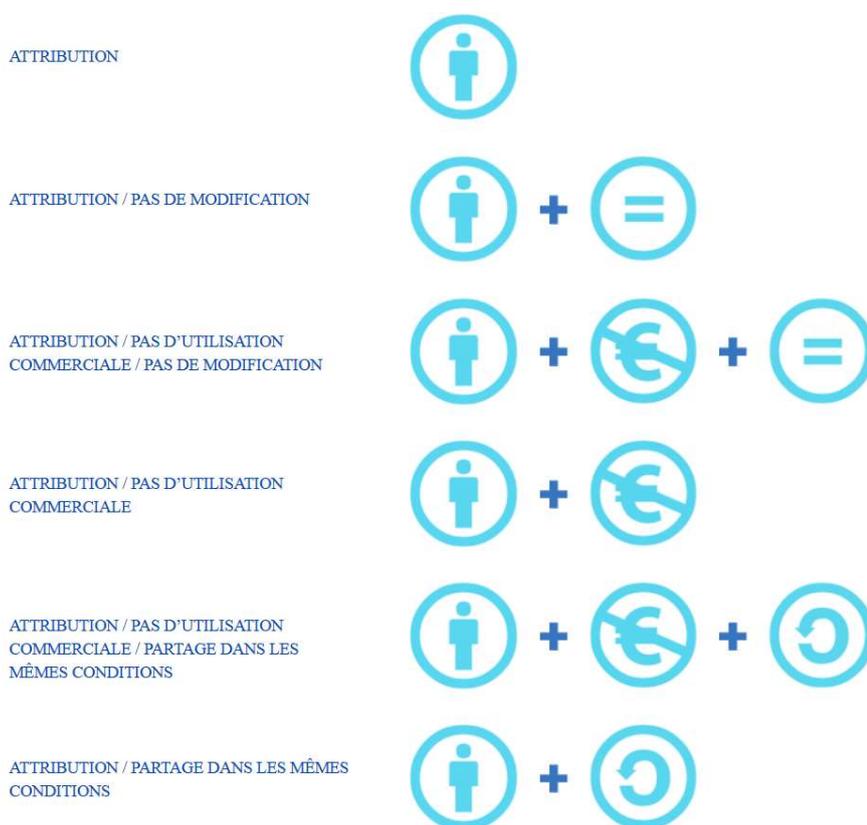


Rediffusion et partage autorisés mais à l'identique.



Interdiction de retirer un bénéfice économique de la réutilisation ou du partage de l'œuvre.

Les différentes combinaisons possibles :



Les quatre options peuvent être arrangées pour créer ces six licences différentes, les six licences CreativeCommons.

Il est courant de vouloir illustrer ses travaux avec une ou des images. On peut donc aller piocher dans du contenu libre, ce qui garantit de ne pas être en infraction avec la législation sur le droit d'auteur.

Des moteurs de recherche permettent de trouver ce type de ressources :

- généraliste (type Google image) avec option dans recherche avancée (1)
- spécialisé <http://search.creativecommons.org> (ex. : Pixabay pour images libres de droits)

Il est possible de chercher en 5 langues différentes, du contenu multimédia, texte, photos, musique et vidéo sur diverses plateformes : Google, Yahoo, Flickr, Jamendo, etc.

Approfondissement : Lancer la vidéo explicative : <https://www.20minutes.fr/high-tech/2246663-20180402-video-fact-checking-trois-conseils-identifier-image-internet>

(1) Dans Google image, après avoir obtenu vos résultats, choisissez « Outils » puis « Droits d'usage » ; dans Bing images, choisissez « Filtre » puis « Licence » ; etc.

Bibliothèque universitaire



Si vous avez des questions :

bib.univ@universite-paris-saclay.fr