

Aspirer un site Web

Pour rédiger un rapport ou bien pour compléter un dossier, vous avez besoin d'informations provenant de pages sur Internet. Or, à moins de disposer en permanence d'une connexion à Internet, y compris en déplacement, ou bien d'avoir pris le temps de récupérer les informations qui vous sont nécessaires avant de bouger, vous n'aurez pas matière à travailler.



La solution : copier sur son ordinateur les pages ou les sites Web qui vous intéressent pour les emporter avec vous et pouvoir ainsi les consulter hors connexion. Cela vous permettra également de **conserver une copie sur votre ordinateur de pages intéressantes**, qui peuvent très bien changer ou disparaître sur Internet.

Les aspirateurs de sites Web sortis en même temps que la démocratisation d'Internet, pour pouvoir télécharger un site en bas débit et le consulter librement ensuite, sont toujours d'actualité. Avec eux, il est possible de **demande le téléchargement de l'intégralité d'un site**, ou seulement d'un certain nombre de niveaux de profondeur des pages dans l'arborescence, de **filtrer les éléments à ramener** : pages, images, animations, vidéo, musique, et même de limiter le poids total du site en fixant un seuil à ne pas dépasser. Des fonctionnalités que l'on retrouve dans HTTrack, un logiciel gratuit.

Et pour ne pas avoir à rapatrier l'intégralité des pages lors d'une mise à jour d'un site que vous avez aspiré, **seules les pages modifiées sont téléchargées**.

Il vous restera alors plus qu'à copier les fichiers sur votre ordinateur portable ou sur une clé USB pour pouvoir naviguer en déplacement sur les pages que vous avez téléchargées, **comme si vous étiez connecté**.

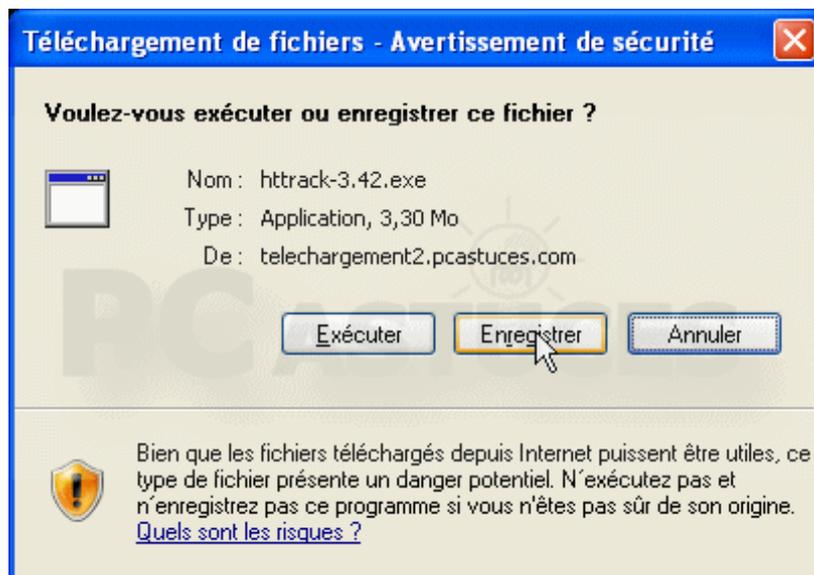
Installer HTTrack

HTTrack est un aspirateur de sites Web complet et gratuit.

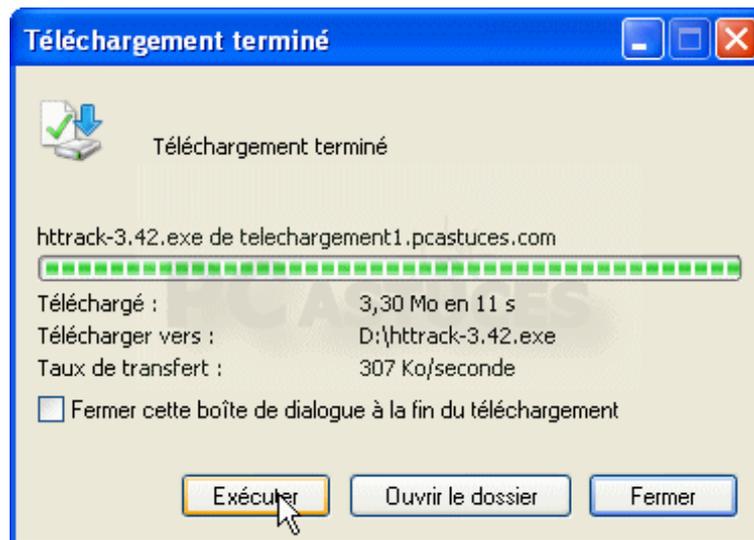
1. Rendez-vous sur [la fiche de HTTrack dans la logithèque](#).
2. Cliquez sur le lien **Télécharger gratuitement**.



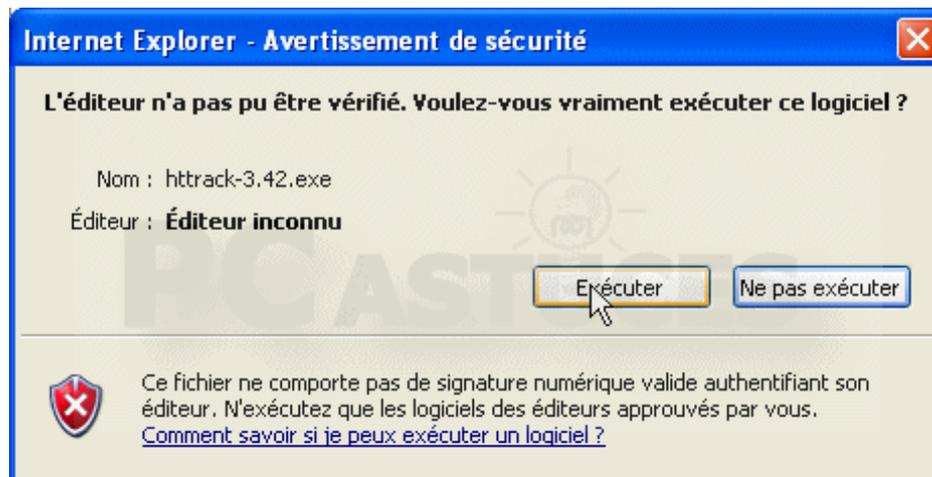
3. Cliquez sur le bouton **Enregistrer**, choisissez un dossier où placer le fichier d'installation de HTTrack puis cliquez une nouvelle fois sur le bouton **Enregistrer**.



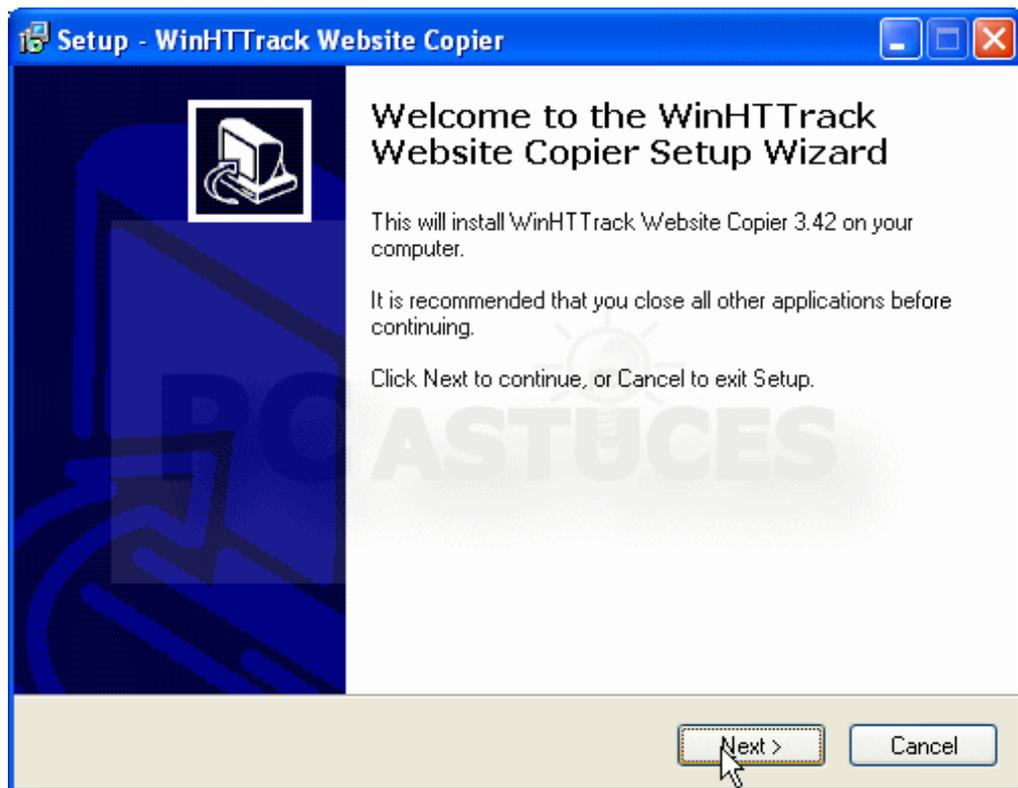
4. Une fois le téléchargement terminé, cliquez sur le bouton **Exécuter**.



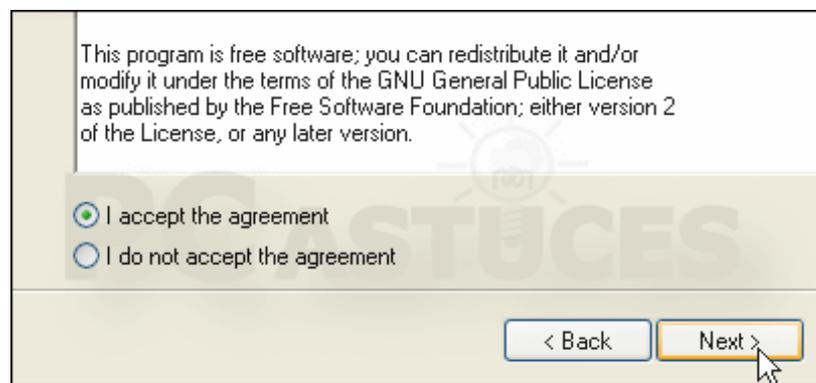
5. Confirmez le lancement du logiciel d'installation en cliquant sur le bouton **Exécuter**.



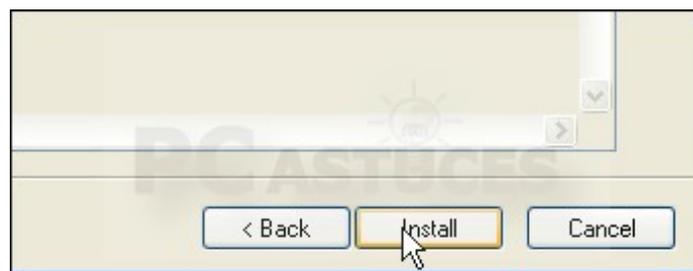
6. L'assistant d'installation de HTTrack s'ouvre alors. Cliquez sur le bouton **Next**.



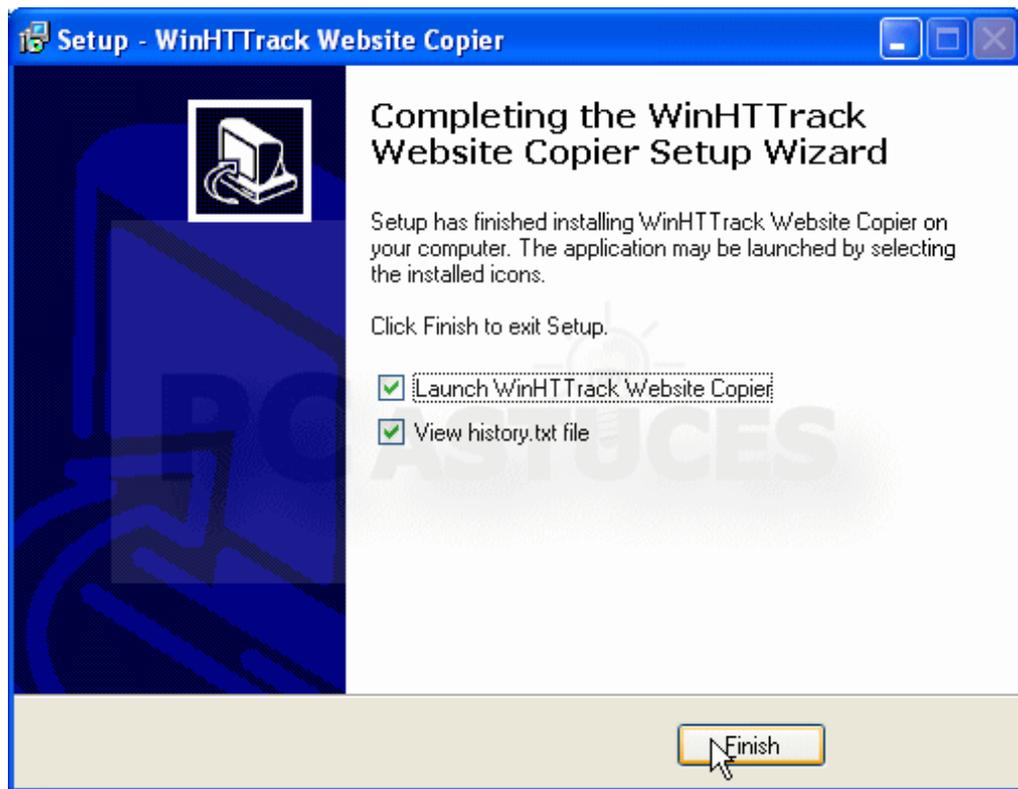
7. Sélectionnez l'option **I accept the agreement** puis cliquez quatre fois sur le bouton **Next**.



8. Cliquez alors sur le bouton **Finish**.



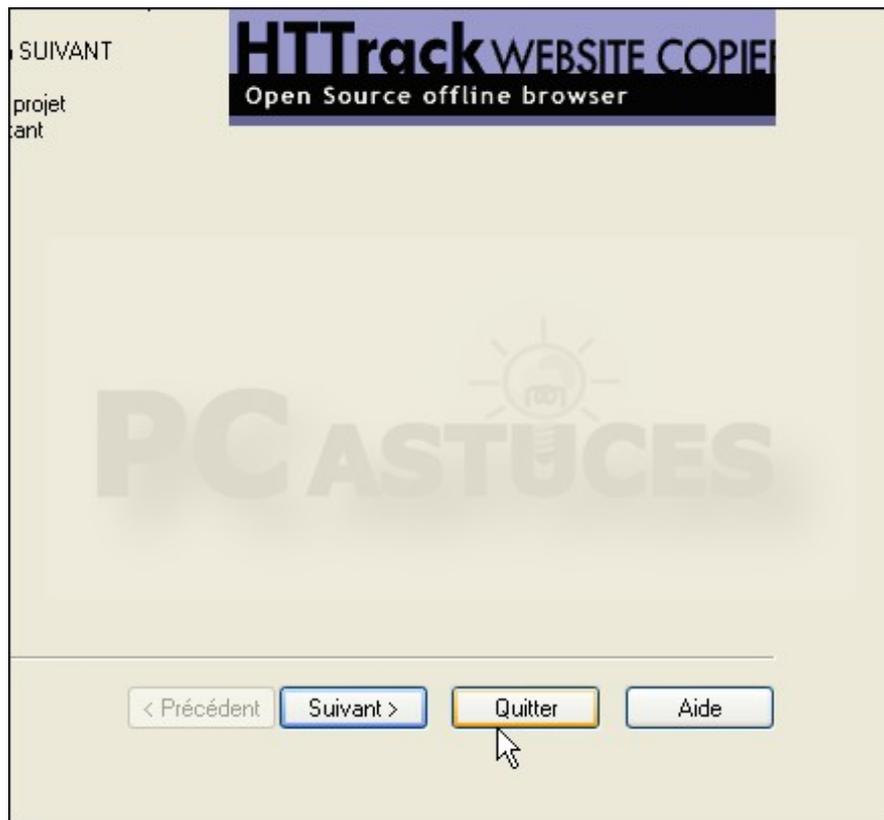
9. Une fois la copie des fichiers terminée, cliquez sur le bouton **Finish**.



10. HTTrack démarre alors. Déroulez la liste **Language preference** puis sélectionnez l'option **Français**. Cliquez sur le bouton **OK**.



11. Fermez alors HTTrack en cliquant sur le bouton **Quitter**.



12. Redémarrez alors le logiciel en double cliquant sur son icône sur le Bureau. L'interface est alors entièrement en français.

Démarrer une capture de site

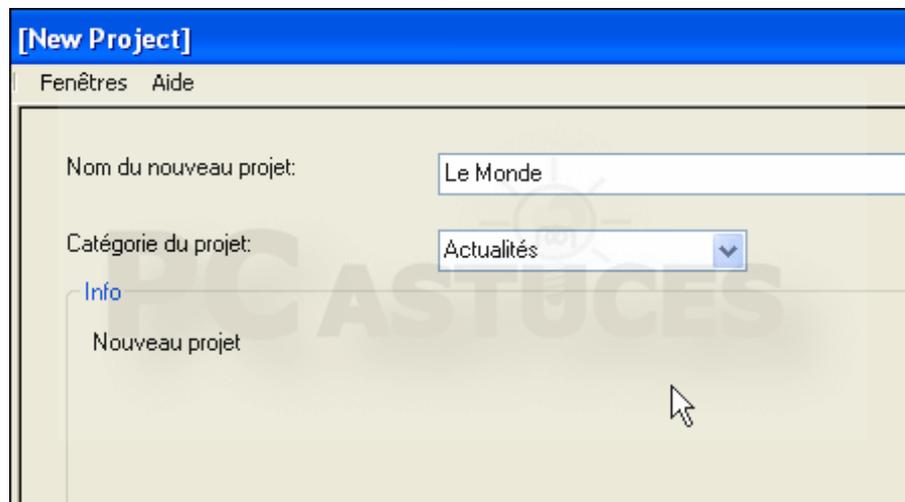
1. Lancez HTTrack.



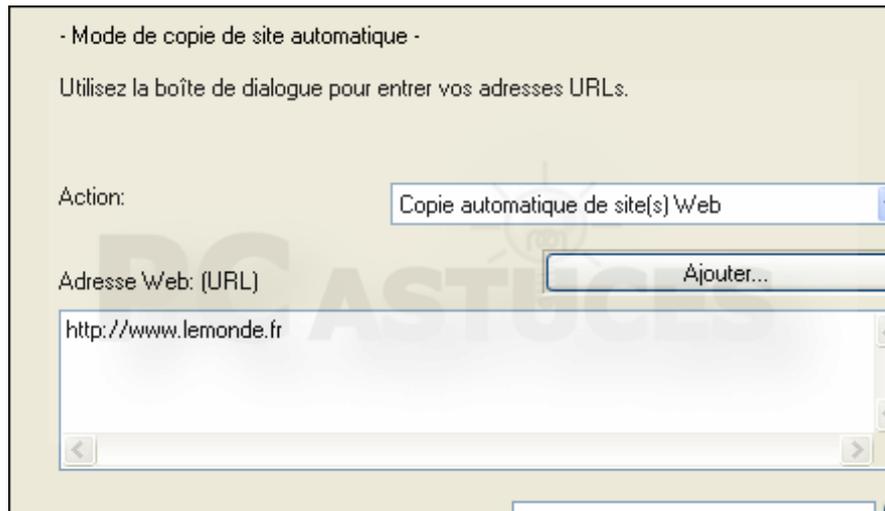
2. Cliquez sur le bouton **Suivant**.



3. Donnez un nom à votre projet de capture puis définissez si besoin une catégorie.



4. Cliquez ensuite sur **Suivant**.
5. Dans le champ **Adresse Web : (URL)**, saisissez l'adresse du site ou de la page que vous souhaitez télécharger. Pour télécharger plusieurs adresses, saisissez-les en les séparant par un retour à la ligne.

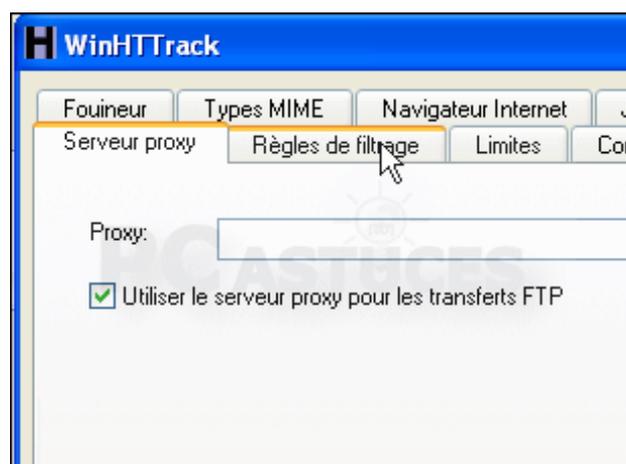


Paramétrer la capture

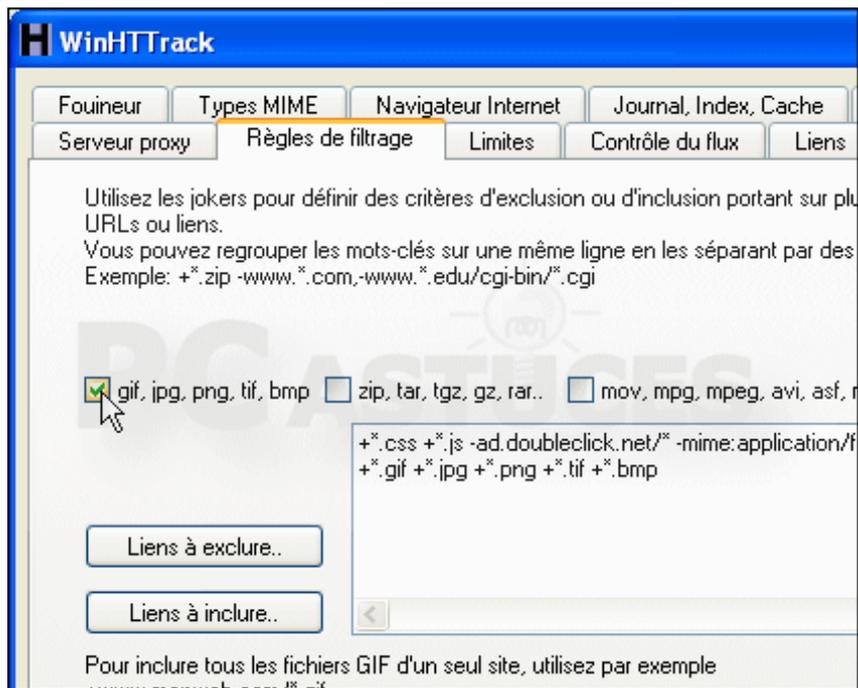
1. Cliquez ensuite sur le bouton **Définir les options**.



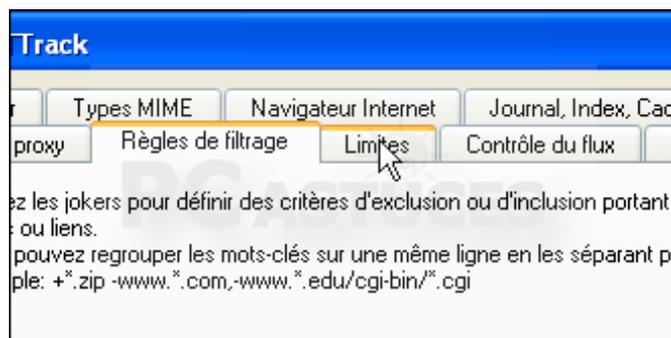
2. Ouvrez l'onglet **Règles de filtrage**.



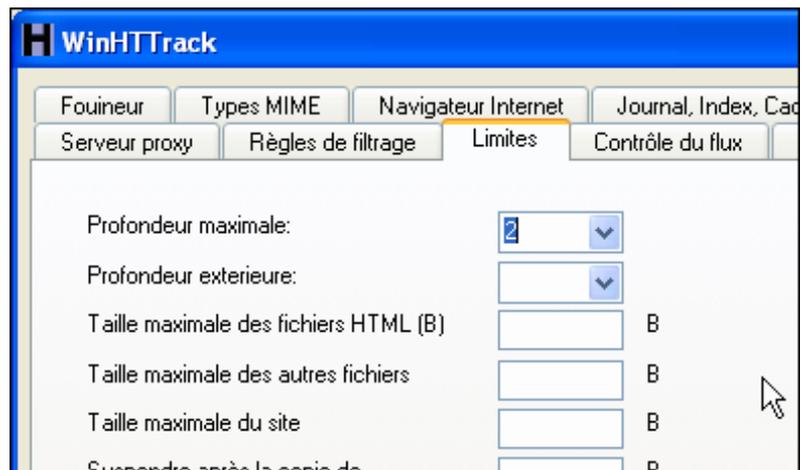
3. Cochez la case **gif, jpg, png, tif et bmp** pour télécharger les images des pages ramenées. Cochez les cases adéquates pour téléchargement aussi les musiques, animations, etc.



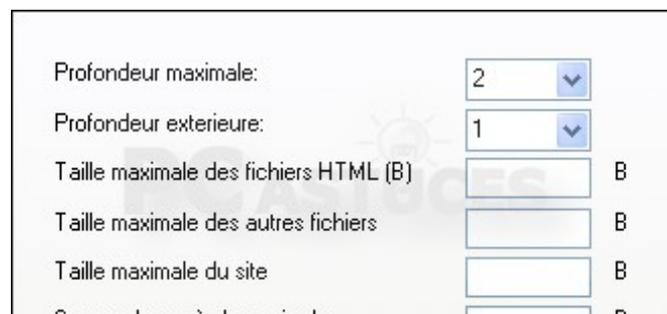
4. Ouvrez ensuite l'onglet **Limites**.



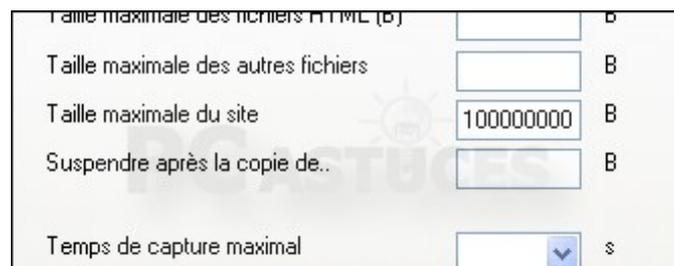
5. Vous pouvez maintenant définir le niveau de profondeur maximal que vous allez télécharger dans l'arborescence de la page que vous avez indiquée. Le niveau **1** correspond à la page d'accueil, **2** à la page d'accueil et à tous les liens qu'elle contient, etc ... Attention à ne pas mettre un niveau trop élevé (au dessus de 3 par exemple) car le nombre de pages téléchargées peut rapidement atteindre la dizaine de milliers sur un site bien fourni.



6. Spécifiez également le niveau de téléchargement des liens qui pointent vers des pages extérieures au site d'origine, **1** par exemple.



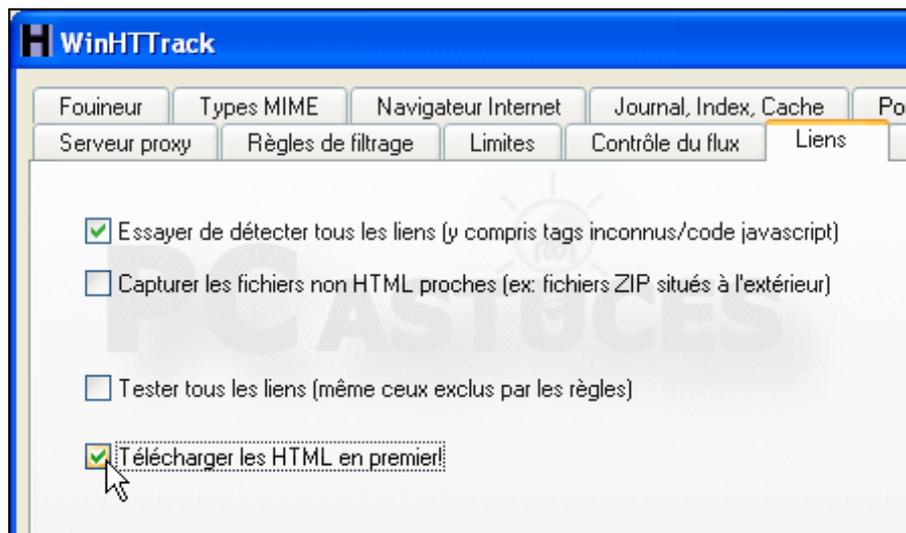
7. Si vous êtes limité en taille, vous pouvez définir une taille maximale du site en octets, **100000000** pour 100 Mo par exemple.



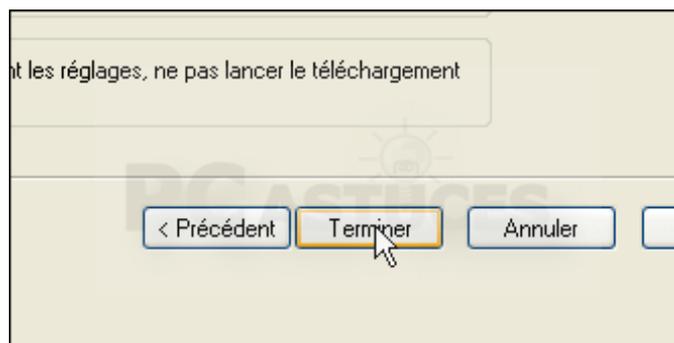
8. Déroulez la liste **Taux maximal** et choisissez l'option **50 000** pour augmenter la vitesse de téléchargement des pages et images.

Temps de capture maximal	<input type="text"/>	s
Taux maximal	<input type="text" value="50000"/>	B/s
Nb. max connexions / secondes	<input type="text"/>	cnx/s

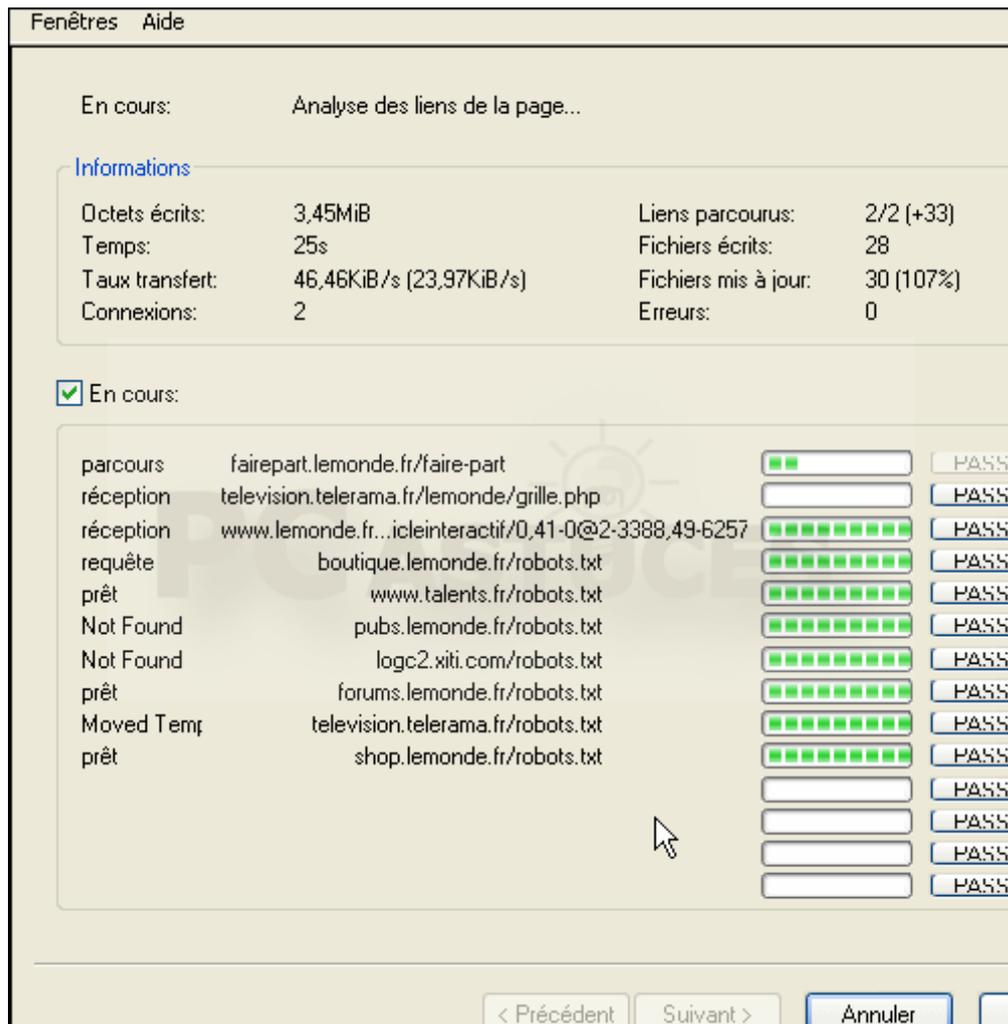
9. Ouvrez ensuite l'onglet **Liens**. Cochez la case **Téléchargement les HTML en premier**. Toutes les pages Web seront ainsi téléchargées avant les images.



10. Cliquez enfin sur le bouton **OK**. Cliquez sur Suivant puis sur **Terminer**.



11. La capture commence alors. Attention, l'opération peut être assez longue : tous les liens sont analysés, les images et pages téléchargées et l'architecture du site recrée sur votre [disque dur](#).



Naviguer sur le site hors [connexion](#)

Toutes les pages et images du site aspiré (dans la limite du niveau que vous avez défini) se trouvent sur votre ordinateur. Vous pouvez consulter quand vous le souhaitez ces pages, les copier sur une [clé USB](#) pour les emporter avec vous, etc.

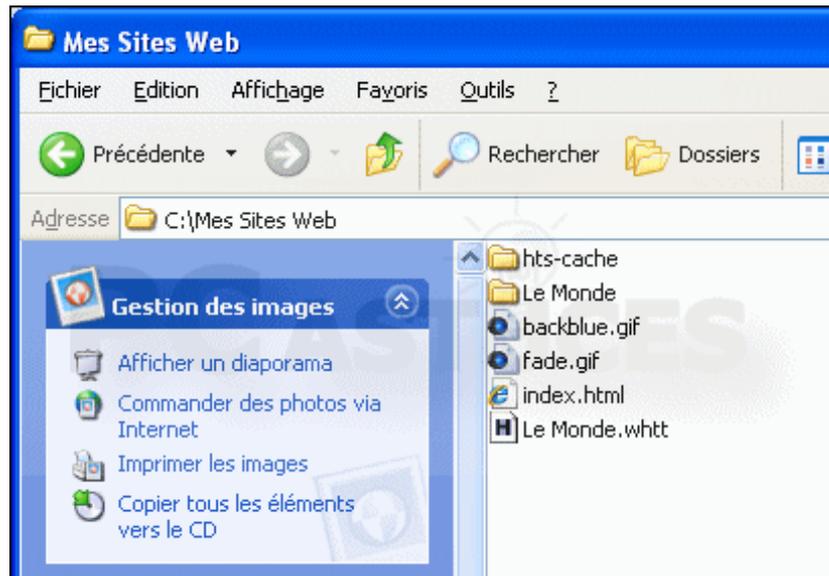
1. Une fois l'aspiration terminée, cliquez sur le bouton **Explorer la copie du site**.



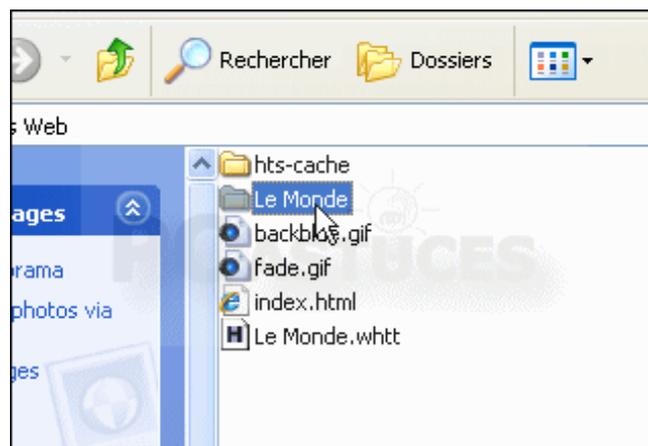
2. Vous pouvez alors naviguer sur le site, hors connexion, directement à partir des fichiers téléchargés.



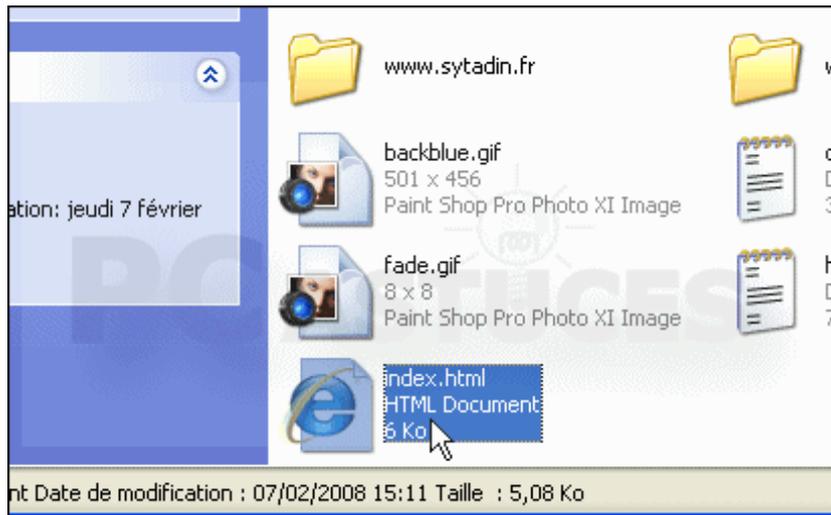
3. Pour emmener le site avec vous, rendez-vous dans le dossier C:\Mes Sites Web\.



4. Copiez le dossier portant le nom de votre projet sur votre clé [USB](#) par exemple.



5. Pour visitez le site, il vous suffira d'ouvrir le dossier et de double cliquer sur le fichier **index.html**.



6. Votre navigateur s'ouvrira alors sur la copie locale du site.



7. Vous pouvez naviguer dedans, comme si vous étiez connecté, dans la limite de la profondeur que vous avez définie.



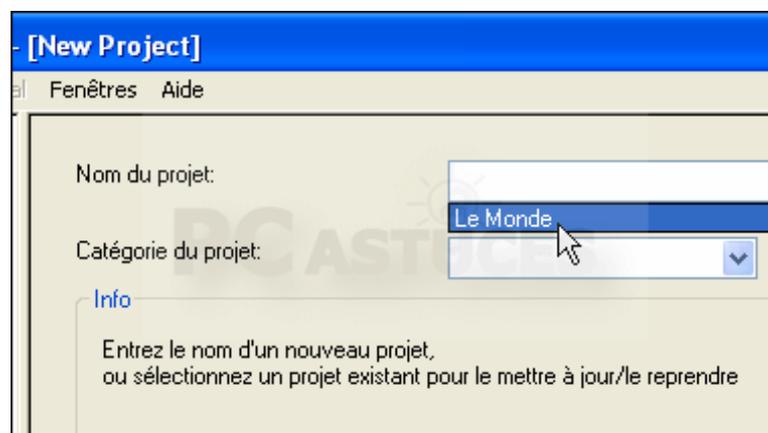
Mettre à jour une capture

Si les pages du site que vous avez aspiré ont été mises à jour, pas besoin de tout ramener. Vous pouvez télécharger uniquement les pages qui ont été modifiées.

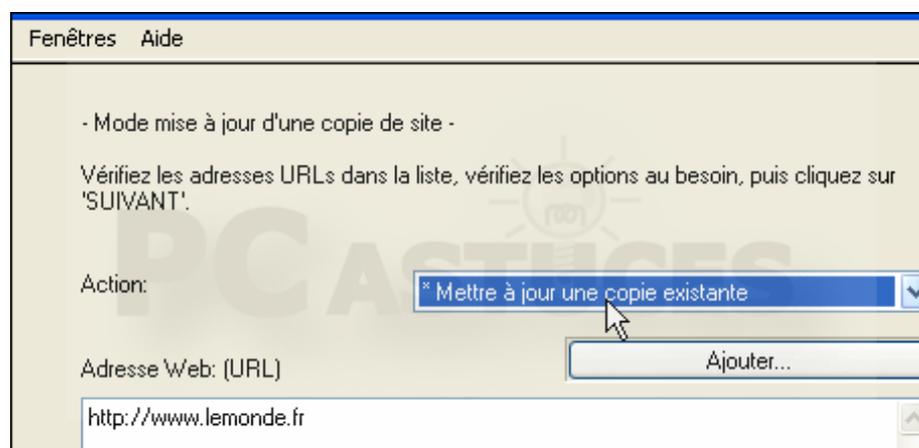
1. Recommencez l'opération depuis l'étape [Démarrer une capture de site](#).



2. Dans le champ **Nom de projet**, déroulez la liste et sélectionnez le nom de votre projet précédent.



3. Cliquez sur **Suivant**.
4. Dans la liste **Action**, vérifiez que l'option **Mettre à jour une copie existante** est bien sélectionnée.



5. Modifiez si besoin la profondeur de téléchargement ou les autres réglages en cliquant sur le bouton **Définir les options**.
6. Cliquez enfin sur le bouton **Suivant** puis sur **Terminer**.

