

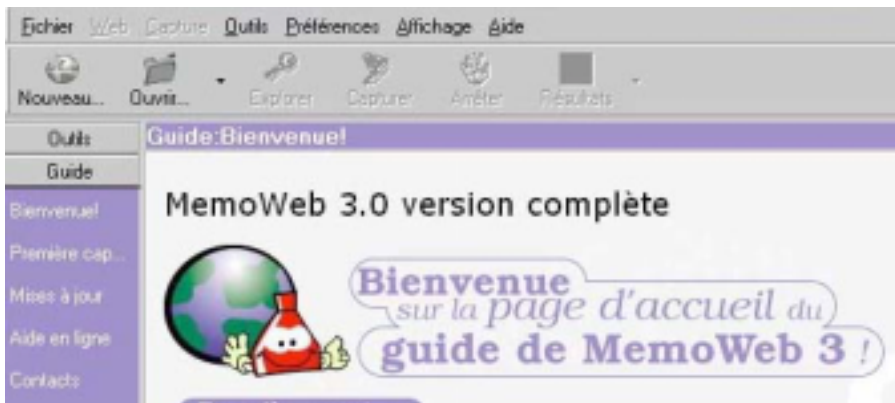
Aspirer un site Internet

Le titre peut paraître mystérieux; on peut dire aussi capturer un site Internet au lieu d'aspirer. Aspirer un site Internet, c'est la possibilité de transférer un site sur son disque dur pour pouvoir ensuite le parcourir en local, c'est à dire sans les problèmes de temps de réponse liés à la connexion. Nous allons examiner 2 logiciels de captures, un sur Windows, MémoWeb 3 de la société nordiste GoTo, l'autre WebDevil de la société américaine Chaotic Software. La version française de WebDevil est le travail du célèbre Jean Pierre Kuypers de l'Université Catholique de Louvain.



MemoWeb 3

Un double clic sur l'icône du logiciel et vous êtes accueilli par un premier écran qui vous souhaite la bienvenue et qui vous propose de vous guider pour votre première capture.

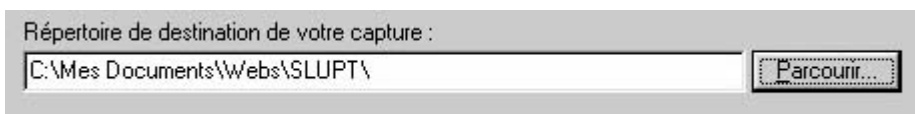


Pour commencer cliquez sur Nouveau, ou laissez vous guider par l'aide qui est proposée. Vous pouvez aussi utiliser une configuration type ou approfondir le paramétrage de la capture .

Ici nous allons capturer l'ensemble du site du SLUPT.

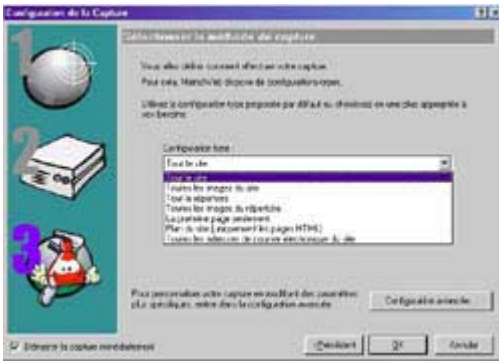


URL est <http://www.slupt.org>



Il faut ensuite choisir un emplacement pour stocker le site sur votre disque dur.

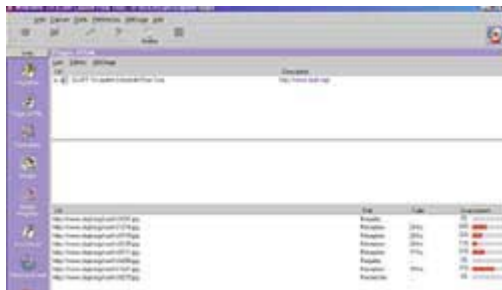
Aspirateur



Configuration-type :

- Tout le site
- Tout le site
- Toutes les images du site
- Tout le répertoire
- Toutes les images du répertoire
- La première page seulement
- Plan du site (uniquement les pages HTML)
- Toutes les adresses de courrier électronique du site

Maintenant nous allons choisir le type de chargement. Ici nous allons télécharger le site complet.



MemoWeb lance plusieurs requêtes simultanées. La structure du site sera recrée sur votre disque local; vous pourrez mettre à jour ultérieurement le site capturé, en ne téléchargeant que les pages modifiées.

MémoWeb ajoutera un logo sur chaque page capturée.

http://www.slupt.org/conf/c0201.jpg	Requête	...	0%	████████████████████
http://www.slupt.org/conf/c1214.jpg	Réception	24 Ko	43%	██████████████████
http://www.slupt.org/conf/c0118.jpg	Réception	20 Ko	32%	██████████████████
http://www.slupt.org/conf/c0125.jpg	Réception	20 Ko	12%	██████████████████
http://www.slupt.org/conf/c0111.jpg	Réception	17 Ko	37%	██████████████████
http://www.slupt.org/conf/c0208.jpg	Requête	...	0%	████████████████████
http://www.slupt.org/conf/c1221.jpg	Réception	15 Ko	72%	██████████████████
http://www.slupt.org/conf/c0215.jpg	Recherche	...	0%	████████████████████

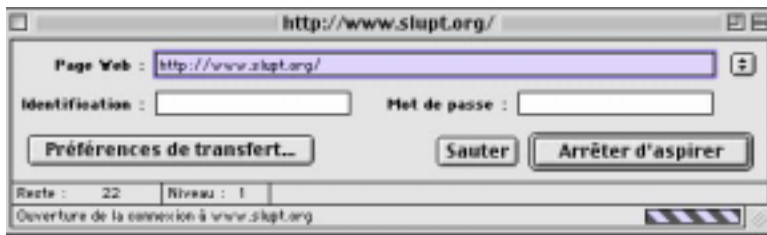


Saint Laurent Université Pour Tous
www.slupt.org



Après avoir lancé le programme, allez dans fichier Nouveau Web Devil puis tapez l'URL que vous voulez capturer.

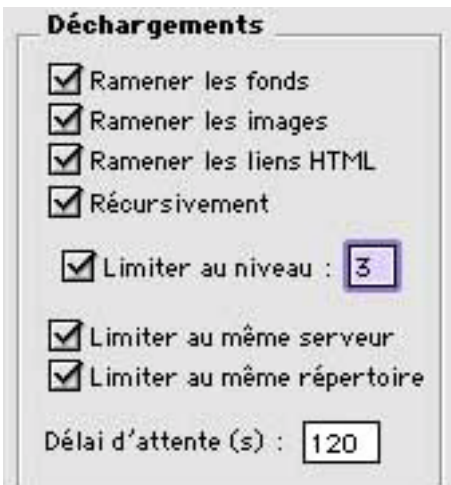
WebDevil est un partagiciel (shareware) commercialisé par Chaotic Software sur Internet au prix de 10 dollars pour la mise à jour depuis une version précédente, ou 3\$ dollars. Le paiement se fait par le système sécurisé par Kagi avec cryptage SSL.



Le paramétrage se fait à partir des préférences.

Dans les préférences vous y indiquerez les éléments à ramener (fonds, images ...), le nombre de niveau dans l'arborescence du serveur et surtout la limite au même serveur.

Vous pourrez aussi mettre à jour sur votre disque par rapport à une précédente capture de ce site.



Attention si vous ne limitez pas au même serveur, vous risqueriez de capturer des milliers de sites. De même, si l'adresse est de la forme sous URL/répertoire vous aurez certainement à limiter au même répertoire.

