

Sieciowi niewolnicy

Wędrowka po światowej pajęczynie, bez psującego zabawę tykania licznika impulsów Telekomunikacji Polskiej S.A. w tle, to marzenie wielu internautów korzystających z modemów. Można je po części zrealizować za pomocą robotów sieciowych – specjalizowanych programów, które pozwalają zminimalizować koszty sieciowych eskapad.

Zadaniem tego typu programów jest zapisanie na lokalnym dysku użytkownika zawartości serwerów WWW w taki sposób, aby możliwe było ich późniejsze przeglądanie bez fizycznego połączenia z Siecią. Ze względu na sposób działania podzielić je można na kilka grup. Pierwszą tworzą *FreeLoader*, *NearSite*, *WebWhacker* i *Net-Attaché Pro*. Programy te oferują bardzo bogate funkcje zarządzania zdobytą informacją. Umożliwiają komfortowe śledzenie bardzo dużej liczby serwisów sieciowych. Są przeznaczone dla osób, które chcą mieć jednoczesny dostęp do wielu źródeł informacji. Nieco inne zastosowanie mają programy z drugiej grupy: *Tierra Highlights* i wzmiankowany już *NetAttaché*. Są adresowane do sieciowych łowców sensacji – dzięki opcji porównywania poprzedniej i obecnej zawartości wskazanej strony natychmiast sygnalizują użytkownikowi nie tylko fakt pojawienia się nowości, ale do-

datkowo zaznaczają zmienione fragmenty stron zwalniając go z konieczności „ręcznego” wyszukiwania tych zmian. Oczywiście zapisują także strony lokalnie umożliwiając ich przeglądanie bez fizycznego połączenia z Siecią. *Teleport Pro* natomiast jest przedstawicielem innej grupy programów, oferujących znacznie skromniejsze instrumentarium (nie posiadających mechanizmów zarządzania danymi pochodzącymi z wielu serwisów), ale doskonale rekompensujących tę niedogodność całkowitą przenośnością zdobytych danych. O ile bowiem wszystkie inne programy zapisują odczytane strony we własnym formacie, o tyle *Teleport Pro* potrafi stworzyć wierną kopię sieciowego serwisu, łącznie z odwzorowaniem struktury katalogów i zachowaniem pierwotnych nazw plików. Wyniki pracy *Teleporta* można kopiować i przenosić zapominając o aplikacji, która umożliwiła ich zdobycie.

Zgromadzone na dysku strony można wyświetlić przy użyciu tradycyjnych przeglądarek WWW (na przykład Netscape Navigator czy Internet Explorera).

Wszystkie produkty za wyjątkiem *Tierra Highlights* umożliwiają określenie maksymalnej ilości miejsca na dysku, jakie chcemy przeznaczyć na informacje ściągane z wybranego miejsca Pajęczyny. Pozwalają także ustalić, ile poziomów hierarchii danego serwera chcemy załadować oraz czy mają być odczytywane strony pochodzące z innych miejsc. We wszystkich programach poza *WebWhackerem* i *Tierra Highlights* możliwe jest też ograniczenie czasu trwania procesu ściągania. Wszystkie programy dysponują terminarzem (scheduler), za pomocą którego określić można częstotliwość „odświeżania” lokalnej kopii serwisu. Zwykle robota można zaprogramować tak, by w nocy samoczynnie łączył się z Internetem, ściągał nowe lub zmodyfikowane strony.

Za wyjątkiem *FreeLoadera*, prezentowane roboty potrafią pomijać podczas ściągania określone typy plików lub zawęzić je tylko do jednego, wybranego typu (np. obrazków).

Żaden program nie przewiduje możliwości świadomego wyłączenia z procesu ładowania poszczególnych drzew katalogowych danego serwisu WWW.

Gdy najważniejsza jest informacja...

Podstawową funkcją większości robotów sieciowych jest dostarczenie użytkownikowi najnowszej wersji wybranych przez niego serwisów. Są one zoptymalizowane ► 93



pod względem funkcjonalności – ściąga-
ne z Sieci dane konwertują z języka
HTML do postaci, która pozwala sprawnie
realizować wyszukiwanie fraz bądź ła-
two zarządzać zdobytymi informacjami.
Podejście takie ma w wielu przypadkach
kluczowe znaczenie, gdyż dzięki temu
uzyskujemy wysoką funkcjonalność pro-
gramów (szybkie wyszukiwanie, oszczęd-
ność miejsca na dysku). Utrudnia jednak,
a czasami nawet uniemożliwia przeniesie-
nie danych na inny komputer.

FREELoader 2.5



Jedynym testowanym produktem, z któ-
rego można korzystać bezpłatnie jest Free-
Loader. Nawiąsem mówiąc, nie jest on
zupełnie bezpłatny – aby móc się nim po-
sługiwać, trzeba z „dobrodrojstwem in-
wentarza” przejść sporą liczbę reklam
sponsorów, którzy umożliwili jego po-
wstanie i dystrybucję. Reklamy te (a jest
ich ok. 6 MB) odczytywane są z Sieci na
początku instalacji. Przy połączeniu mo-

dany serwis ma być przeglądany w poszu-
kiwaniu zmian. Gdyby dostęp do danego
miejsca chroniony był hasłem, można je
wprowadzić w odpowiednie pola
dialogowe.

Łałowana w ten sposób zawartość
wybranego serwisu możemy przeglądać
zwykłą przeglądarką w trybie offline.
Niestety, FreeLoader nie dysponuje
funkcją wyszukiwania ułatwiającą od-
najdowanie haseł w dokumentach ścią-
gniętych z Sieci.

NEARSite

Po uruchomieniu NearSite, na ekranie
ukazuje się niewielki pasek narzędziowy
z siedmioma przyciskami. Za ich pomocą
można otworzyć okno z zakładkami (ad-
resami interesujących nas miejsc) lub uru-
chomić zintegrowany program *Cache
Explorer* (pozwalający przeglądać ścią-
gnięte już strony). Możliwe jest także
uruchomienie przeglądarki WWW
i oczywiście wywołanie funkcji **Auto-
browse** powodującej łałowanie na
dysk twardy wybranych miejsc.

NearSite wykazuje szczególną ela-
styczność, gdy chodzi o ograniczenie ilo-
ści ściąganych danych. Możliwe jest
określenie maksymalnego czasu trwania
ściągnięcia oraz wielkości miejsca na dys-
ku twardym przeznaczonego na dane.
Co więcej, z procesu ładowania wyłą-
czyć można grafikę, aplety Javy, wszyst-
kie osadzone w stronach obiekty, a tak-
że pliki, których rozmiar przekracza
określoną wielkość. Program zapamię-
tuje ściągnięte dane w swym własnym
formacie, dlatego można je przeglądać
tylko za pomocą wbudowanego modułu
Cache Explorer. Dysponuje on rozbudo-
wanymi funkcjami wyszukiwania i selek-

Przeglądarki offline – dla kogo?

Za sieciowymi robotami przemawia przede
wszystkim to, iż pozwalają one użytkowni-
kom oszczędzić na opłatach za korzysta-
nie z Sieci. Sprawdźmy: za około 150 zło-
tych, które większość tych programów ko-
sztuje, można surfować po Internecie
około 40 godzin.

Ktoś kto potrzebuje jedynie pojedyn-
czych stron, może je równie dobrze zapisać
na dysku za pomocą przeglądarki jako pliki
HTML i następnie otwierać je później wielo-
krotnie. Osadzoną grafikę trzeba przy tym
ściągać oddzielnie – chyba, że możemy się
bez niej obyć. Jedynie ktoś, kto regularnie
korzysta z większej liczby serwisów i prze-
gląda przy tym wiele stron WWW, a potrze-
bne mu informacje chce archiwizować na
własnym komputerze, może dzięki tego ty-
pu programom zaoszczędzić w dłuższej
perspektywie trochę pieniędzy.

Za sieciowymi robotami przemawiają też
i inne względy: ponieważ strony w Sieci
zmieniają się bardzo często, to lokalne ar-
chiwizowanie ważnych danych z jakiegoś
miejsca może okazać się bardzo przydatne.
Z takiego archiwum korzystać można na-
wet wtedy, gdy oryginału dawno już nie ma
w Sieci. W przypadku wielkiej ilości danych,
nieocenione usługi przy odnajdywaniu po-
trzebnych informacji może oddać rozbudo-
wana funkcja wyszukiwania, którą oferują
wszystkie (poza FreeLoaderem i Telepor-
tem Pro) opisywane programy.

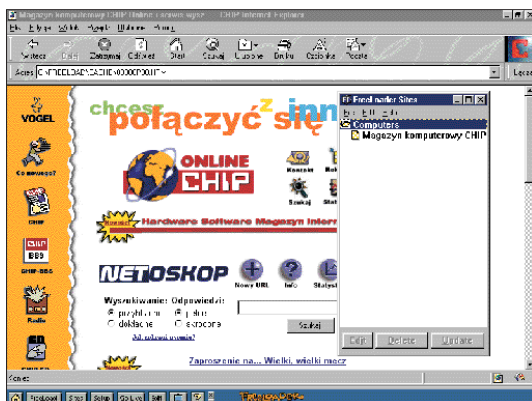
cji łałowanych plików. Pozwala także
na zarządzanie zapamiętanymi lokalnie
serwerami – może wyświetlić wszystkie
należące do jakiegoś serwisu strony, po-
sortować je według adresów lub daty
ostatniej modyfikacji, a także wykaso-
wać te, które nie są już potrzebne.

NearSite odczytane z Sieci pliki zapisu-
je w swoim własnym formacie. Na szcze-
ście oferuje jednak możliwość eksportu
danych do formatu HTML, dzięki czemu
strony WWW można również przeglądać
w zwykłej przeglądarce, nie dysponując
programem NearSite.

Sposób współpracy z tym programem
(osobne okna Bookmarks i *Cache Explo-
rer*) wymaga od użytkownika poświęce-
nia pewnego czasu na opanowanie sposo-
bu jego obsługi.

NET ATTACHÉ PRO 2.1

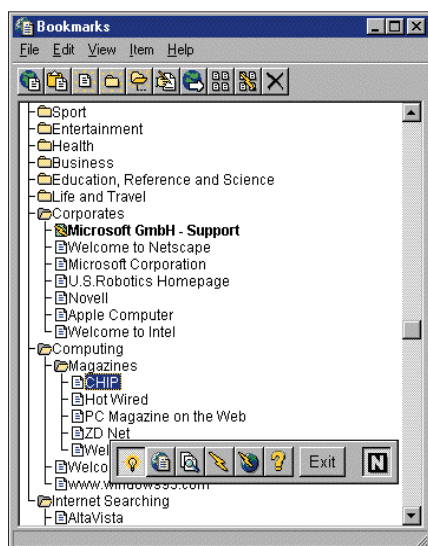
Podobnie jak w pozostałych programach,
serwis przeznaczony do ściągnięcia moż-
na wybrać za pomocą przeglądarki
WWW. NetAttaché używa do klasyfiko-
wania interesujących nas serwisów siecio-
wych tak zwanych „Briefs”, zestawów ad-
resów udostępniających informacje z jed-
nej dziedziny. Oferuje szerokie możliwo-
ści wybierania typów danych, które



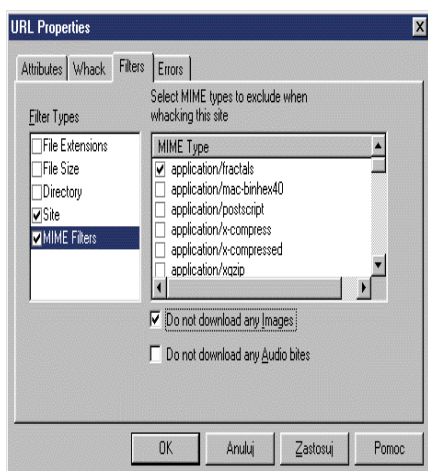
FreeLoader jest całkowicie zintegrowany z przeglądarką WWW

demowym (np. 28,8 KB) trwa to około
45 minut. Na dodatek FreeLoader auto-
matycznie instaluje szereg kanałów,
którymi na bieżąco spływają na przykład
najświeższe wiadomości sportowe. Pro-
gram udostępnia specjalny wygaszacz
ekranu wyświetlający zgromadzone in-
formacje.

FreeLoader współpracuje z Netscape
Navigatorem począwszy od wersji 2.0
i Internet Explorerem od wersji 3.0. Po
zakończeniu instalacji, przy dolnej kra-
wędki okna używanej przeglądarki poja-
wia się listwa z przyciskami. Aby ściągnąć
interesującą nas stronę (strony) otwiera-
my ją przy użyciu przeglądarki WWW
i klikamy przycisk **Freeload**. W poja-
wiającym się okienku określamy paramet-
ry śledzenia zmian – ile poziomów ser-
wisu chcemy łałować, czy przy ściągnięciu
mają być uwzględniane strony znajdujące
się na innych serwerach oraz jak często



Trzeba się przyzwyczaić: Near Site to tylko jeden pasek narzędzi i jedno okno



WebWhacker 3.0 oferuje liczne opcje ograniczające ładowanie

zostaną wyłączone z procesu ładowania (można zrezygnować z grafiki, plików Plug-In, ramek, apletów Javy czy kontrolerek ActiveX).

NetAttaché pozwala załadować do 99 poziomów danego serwisu sieciowego. Chociaż dla każdego miejsca oddzielnie można wybrać, które pliki i ile poziomów będzie uwzględnianych, to możliwość ograniczenia wielkości ściąganych danych okazuje się mniej elastyczna. Obszar dysku twardego, jaki mogą zająć dane, określić można jedynie sumarycznie dla wszystkich serwisów należących do jednej kategorii. Podobnie czas trwania ładowania oraz częstotliwość odświeżania zawartości można ograniczać tylko wspólnie, dla wszystkich miejsc w obrębie jednej kategorii.

Plusem NetAttaché są natomiast filtry. Pozwalają one na przykład wybierać i wyświetlać tylko te strony, które uległy ostatnio zmianie. Zmodyfikowane elementy stron zaznaczane są czerwonym wykrzyknikiem. Za pomocą funkcji *Smart Search* można poszukiwać haseł w tytułach lub tekstach stron i ściągać tylko te, w których one występują.

Funkcja pełnotekstowego wyszukiwania pozwala korzystać z operatorów logicznych (AND, OR, NOT) i nawiasów. Dla globalnych poszukiwań w Sieci NetAttaché oferuje dostęp do Yahoo!, można także dodać inne maszyny informacyjne. Obok kosztującej 50 dolarów wersji Pro NetAttaché dostępna jest także bezpłatna wersja Light. Nie dysponuje ona jednak ani funkcjami wyszukiwania i filtrowania, ani schedulerem pozwalającym zautomatyzować operację ładowania.

WEB WHACKER 3.0

Historycznie – Web Whacker był jednym z pierwszych robotów sieciowych. Najnowsza wersja oferuje już usługi Kreato-

ra, służącego do pierwszego ładowania zawartości wybranego serwisu. Udostępnia także możliwość kategoryzowania interesujących nas serwisów. Odczytane z Sieci strony wyświetlane są w formie drzewa, co znacznie poprawia przejrzystość informacji.

Z procesu ładowania wykluczyć można praktycznie każdy, dowolny typ plików. Opcję tę ustawia się osobno dla poszczególnych adresów, podobnie jak terminarz (scheduler). On jednak pozwala określić jedynie, czy odświeżanie informacji ma się odbywać raz dziennie, raz w tygodniu, czy raz na miesiąc. Czas rozpoczęcia wyszukiwania zmian, ustawia się globalnie dla wszystkich obserwowanych serwisów. Nie jest niestety możliwe także ograniczenie ilości lokalizowanych danych (miejsce na dysku) lub czasu trwania ładowania. Przy pracy w trybie online bardzo praktyczny okazuje się pasek narzędziowy, oferujący dostęp do najważniejszych funkcji WebWhackera w sytuacji, gdy jego okno schowane jest w tle.

Funkcje wyszukiwania, z pomocą których w załadowanych serwisach można odnajdywać zarówno hasła proste, jak i ich kombinacje logiczne, porównywalne są z podobnymi narzędziami NetAttaché i NearSite. WebWhacker dysponuje dodatkowo funkcją, wyświetlającą przegląd stron zmienionych w ciągu ostatnich „x” dni.

Gdzie te zmiany?

No właśnie – w przypadku niektórych serwisów sieciowych informacja o tym, że udostępniają one nowe wiadomości, nawet poparta możliwością natychmiastowego wglądu w tę nową zawartość,

jest niewystarczająca. Czasem bardzo potrzebujemy wiedzieć co się zmieniło. Sprawa nabiera szczególnego znaczenia w momencie, gdy chcemy śledzić zawartość wielu serwisów jednocześnie. W tym celu wykorzystać można NetAttaché Pro i Tierra Highlights. NetAttaché to prawdziwy „kombajn”, ta funkcja jest tylko jedną z wielu oferowanych przez niego możliwości. Highlights natomiast jest programem dedykowanym do śledzenia i zaznaczania zmian zachodzących na stronach.

TIERRA HIGHLIGHTS

Podobnie jak konkurenci pozwala on stworzyć zbiór ulubionych miejsc w Sieci i indywidualnie ustalić częstotliwość ich uaktualniania. Zapisując lokalnie informacje udostępniane przez serwisy sieciowe przede wszystkim porównuje odczytaną właśnie zawartość z tym, co uda-



Tierra Highlights wyraźnie zaznacza zmienione fragmenty stron

ło się „zdobyć” przy poprzedniej aktualizacji. W ten sposób jest w stanie stwierdzić dokładnie, co zmieniło się na każdej

► 96

Przeglądanie pamięci cache w przeglądarce

Przeglądarki WWW tworzą na lokalnym dysku komputera specjalne katalogi (cache) pełniące rolę pamięci podręcznej, w których przechowują ściągnięte z Sieci dane. Jeśli użytkownik naciśnie w przeglądarce przycisk „Back”, to dane potrzebne do wyświetlenia strony zamiast z Sieci pobrane zostaną z pamięci podręcznej. Dane te przechowywane są zwykle w takiej formie, w jakiej zostały odczytane z Sieci. Można je zatem pojedynczo oglądać bezpośrednio w przeglądarce lub za pomocą programu do wyświetlania plików graficznych.

Oprócz samych danych, przeglądarki przechowują także pewne dodatkowe informacje potrzebne do zarządzania zawartością tej pamięci (na przykład dotyczące sposobu powiązania grafiki i dokumentów HTML). Program Cache Explorer (CE) umożliwia komfortowe, wielokrotne przeglądanie

stron znajdujących się w pamięci podręcznej. Korzystając z interfejsu podobnego do Eksploratora Windows, CE wyświetla po lewej stronie adres ostatnio otwieranych w przeglądarce miejsc, po prawej – pochodzące z nich pliki. Dwukrotne kliknięcie ikonki pliku HTML w prawym oknie otwiera wskazany dokument w przeglądarce WWW. CE oferuje także bardzo wygodną funkcję wyszukiwania zadanego tekstu we wszystkich dostępnych bądź tylko pochodzących z określonego serwera dokumentach.

Różnica pomiędzy tym programem, a innymi prezentowanymi tu produktami polega na tym, że CE nie zapisuje danych na dysku, a jedynie umożliwia wyświetlanie tych plików, które już znajdują się w cache'u przeglądarki. W chwili, gdy przeglądarka usunie z bufora starsze pliki, staną się one niedostępne także dla Cache Explorera.



Zawsze na bieżąco

Pod adresem <http://www.netmind.com/> znaleźć można tak zwany URL-Minder, który wysłanie do nas list, gdy we wskazanym przez nas serwisie sieciowym pojawią się jakieś zmiany. W ten sposób można uniknąć osobistego śledzenia pojawiających się nieregularnie nowości.

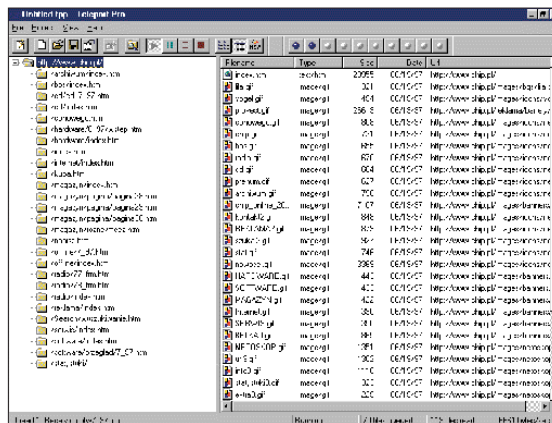
wybranej przez użytkownika stronie. Wyświetlając później zaktualizowaną stronę zaznacza jej zmienione lub nowe fragmenty w wybrany przez użytkownika sposób (wybrany kolorem i/lub zakreślając je dodatkową ramką). Jest to proste, małe, ale szalenie użyteczne narzędzie szybko i przejrzysto informujące o zmianach.

Ciekawą funkcją tego programu jest także możliwość pomijania niektórych zmian zachodzących na stronach. Tierra Highlights pozwala zdefiniować tzw. filtry – zestawy słów, których pojawienie się na danej stronie jest konieczne, aby program zasygnalizował zmianę jej wartości. Filtry pozwalają także opisać zmiany nieistotne dla użytkownika, których obecność ma pozostać „niezauważona” przez program.

TELEPORT PRO 1.28



W porównaniu do opisanych do tej pory programów Teleportowi wyraźnie brakuje swoistej nadbudówki, pozwalającej „wycisnąć” ze zdobytych danych maksimum informacji. Ten program jest po prostu doskonałym narzędziem umożliwiającym łatwe i szybkie stworzenie kopii serwera sieciowego na lokalnym dysku. Zdobyte dane Teleport zapisuje w takiej postaci w jakiej znajdują się one w Sieci. Użytkownik może dodatkowo zdecydować, czy chce lokalnie odtworzyć także strukturę katalogów zdalnego serwera. Ciekawą funkcją udostępnianą przez ten program jest możliwość konwersji nazw odczytywanych plików do postaci 8 + 3 (8 znaków nazwy i 3 rozszerzenia, konwencja DOS-a). Teleport jest jednym z najszybszych robotów sieciowych: potrafi jednocześnie odczytywać z Sieci kilka plików, uruchamiając dla każdego z nich osobny proces. Wszystko to dzieje się automatycznie, a liczba równoległo pracujących robotów dopasowywana jest do przepustowości wykorzystywanego łącza i wydajności zdalnego serwera. Unikatową funkcją Teleporta jest jego umiejętność eksplorowania tzw. server-side maps. Równie ciekawa jest możliwość odтворzenia jedynie struktury zdalnego serwera bez żmudnego odczytywania danych (ściągane są jedynie informacje o nazwach i położeniu plików). Za pomocą takiego



„szkicu” można łatwo wybrać interesujące nas pliki i je ściągnąć. Podobnie jak pozostałe programy Teleport udostępnia możliwość okresowego uaktualniania informacji, wyłączenia z procesu ściągnięcia określonych rodzajów (i wielkości) plików i wiele innych „standardowych” opcji.

Czy warto?

Wszystkie przebadane produkty są zupełnie wystarczające, jeśli chodzi o jednorazowe załadowanie stron WWW, ich późniejsze przeglądanie bez połączenia z Siecią i wyszukanie wśród nich żądanych haseł. Dobrze radzą sobie również z aktualizacją danych.

Oddzielnym problemem jest natomiast prezentacja zmian zachodzących na stronach – to zadanie realizują nieco gorzej. Jedynie NetAttaché Pro pozwala wybierać pomiędzy wyświetlaniem adresów zmienionych miejsc lub tylko zmodyfikowanych stron. I tylko NetAttaché Pro i Tierra Highlights używają specjalnych oznaczeń dla zmienionych fragmentów. WebWhacker i NearSite umożliwiają jedynie przeszukiwanie odczytanych stron oraz filtrowanie i sortowanie ich według daty modyfikacji.

W przyszłych wersjach robotów sieciowych przydałby się graficzny schemat struktury załadowanego miejsca (najlepiej z informacjami, jak duże są poszczególne dokumenty) oraz możliwość wykluczania z ładowania pojedynczych podkatalogów (po części ten postulat spełnia Teleport Pro).

Abstrahując od tych ograniczeń, Tierra Highlights, NearSite i NetAttaché można polecić tym wszystkim użytkownikom, którzy stale odwiedzają pewną większą liczbę miejsc w Internecie i chcieliby być

Teleport Pro wyświetla w postaci drzewiastej strukturę załadowanego miejsca

informowani na bieżąco o zachodzących w nich zmianach. NetAttaché wykazuje pewną przewagę w przypadku dużych serwisów; za NearSite przemawia elastyczność w konfigurowaniu poszczególnych miejsc WWW i rozbudowane funkcje zarządzające. Teleport Pro natomiast to narzędzie dla oszczędnych którzy chcą płacić tylko za czas wykorzystany na transmisję danych lub rzadko mają okazję korzystać z Sieci, a chcieliby w pełni czerpać z jej zasobów.

Piotr Wyrzykowski

Gdzie można zdobyć opisane programy?



FreeLoader 2.5

Freeware

Producent: <http://www.freeloader.com/>

Systemy operacyjne: Windows 3.1x, 95 i NT

NearSite

Pełna 30-dniowa wersja testowa

Cena: 49,95 USD

Producent: <http://www.nearsite.com/>

Systemy operacyjne: Windows 3.1x, 95 i NT

Net Attaché Pro 2.1

Pełna 30-dniowa wersja testowa

Cena: 50 USD (Net Attaché Light dla prywatnych użytkowników jest bezpłatny)

Producent: <http://www.tympani.com/>

Systemy operacyjne: Windows 3.1x, 95 i NT

WebWhacker 3.0

Pełna wersja testowa umożliwi załadowanie 15 stron (tylko 2 poziomy jednego serwisu)

Cena: 49,95 USD

Producent: <http://www.ffg.com/>

Systemy operacyjne: Windows 95 i NT

Cache Explorer 1.25

Pełna 21-dniowa wersja testowa

Cena: 29 USD na przeglądarkę

Producent: http://ourworld.compuserve.com/homepages/M_Wolf/

Systemy operacyjne: Windows 3.1x, 95 i NT

Teleport Pro 1.28

Wersja testowa umożliwia załadowanie 200 dokumentów z jednego miejsca

Cena: 39,95 USD

Producent: <http://www.tenmax.com/>

Systemy operacyjne: Windows 3.1x, 95, i NT

Tierra Highlights

Pełna 30-dniowa wersja testowa

Cena: 49 USD

Producent: <http://www.tierra.com/>

Systemy operacyjne: Windows 3.1x, 95, i NT

Uwaga!

Niektóre opisywane programy znaleźć można na płycie CHIP CD 8/97 w dziale **Software**

| Przeglądarki offline'owe

