

ネットサーファー 御用達、 ホームページの 自動収集ソフト 「波乗野郎」登場！



第2回 自動収集ツール編

illustrator : Kaori Takahashi



指定したホームページの内容を自動的に読み込んでくれるツール「PerMan Surfer 波乗野郎」。

BUG(株)コープ

ども。私が手塩にかけて育てた隊員（おいおい）を編集部に奪われ、前回からまともや1人隊長状態になってしまった村野なのです。

ノッケからなんですけど、考えてみると、この探検隊もかなりの歴史になっちゃいましたですね。ま、当初のISDNからプロバイダー探検、PHSに苦し紛れのWindowsと、まあいろいろありました。これもひとえに皆様の厚いご支援が……って、終わっちゃうわけじゃないんだけど。

それはともかく、最近は苦情のお手紙もいただいてまして、記事が某機種に偏向してるぞ、専用線をつないでいい気になってないか、サーバーはまだか、だんだん読者と距離が出てきているのではないか、どうしていつもそんなにノーテンキでいられるのか（ほっとしてくれ！）……などなど。

こういうご意見が編集会議 担当編集者M嬢 探検隊本部（実はインプレスのサーバー）私、ってな具合にフィードバックされ、そのたびメゲたり、喜んだり。まったく筆者ってやつは……あ、オレか。

そんな事情が錯綜するなか、今回編集部から出されたコマンドは「1人でできるもん挑戦」シリーズ（う～む、どこかで聞いたような）第一弾。ふつーのユーザーがインターネットを使うとき、役にたちそうなグッズ、インターネット生活のための快適・節約・お役立ちモノを取り上げる、という低予算&隊長単独のプロジェクトなのである。

今回の探検ターゲットは「波乗野郎」こと、正式名称「PerMan Surfer 波乗野郎」なんでも北海道にある「BUG」というおそろしげな会社で作ったソフトだそうだが、北海の荒波を見事に乗りこなすスゴ腕な男(?)らしい。

編集部で聞かされたうたい文句では「指定されたURLを巡回し、ディスク上にWebサーバーの内容を先読みしてくれる便利なソフト」。ユーザーは「ディスク上に先読みされたWWWのページをあとでゆっくりと時間を気にしないで読むことができる」というふれこみ。どうして今までこんなものがなかったんだ？のお便利モノだそうだ。

一部方面では、オマケCD-ROMに試用版が入っているというモロ見えの理由から取り上げるようになったというウサがないでもないが、「波乗野郎」と対決するインターネット探検隊。今回は半分だけ初心に戻りつつ、ダイアルアップユーザーのココロ。なつかしの公衆回線とモデムを持ち出して、いざ探検に出発っ~!

7月号CD-ROMの試用版で 波乗りにも挑戦

まずはインターネットマガジンご自慢のオマケCD-ROMをカートリッジに入れてMacintoshに差し込む(うーむ、いまだに古いマシンを使っているのがバレバレ)。

「Mac」と書かれたホルダーが表示されるので、この中を「WWW」「etc」「PerMan Surfer」という具合に開いていって、問題のソフトを取り出す。7月号の付録に入っているのは「version 0.64d6」だ(今月号では0.91d3版を収録)。

ホルダーの中には「PerMan Surfer 68K」と同PPC、そして「naminori」というホルダーが入っている。「はいはい、どうせ私は旧Macですよん」って68Kのアイコンをクリックすることになるのだが、その前に「naminori」の中身を点検だ。

中にはGIFファイルとJPEGファイル、それにHTMLのファイルが……。あれれ、「Readme」はないの?と思ったけど、考えたら「index.htm」をネットスケープで開けばいいの。

というわけで、ネットスケープを立ち上げて「File」メニューから「Open File」で「index.htm」を開く。と、おおっ、画面にはワンちゃんが波乗りしているイラストの「波乗野郎」のホームページ(こういう場合も「ホームページ」って呼ぶ?)が表示されるではないか。ま、そんなに驚くほどのことじゃないけど、波乗野郎の正体はイヌかよ~(と、ちょっと拍子抜け)。

いかにもそれっぽいレイアウトのページを見ると、「What's New」にこのバージョンのリリースデートが書いてある。どれどれ、96年の4月ね。う~ん、ちょっと古いな。で、その下には「0.64d6について」という項目があって「試用期間を1996年6月15日までに延長しました」とある。げ、ひょっとしてこの号が発売されるころには、これって使えないかも、と一瞬アせる。が、ま、その頃になればまた新しいバージョンくらい、なんとかなるだろう、と前向きに考えないとこの手の記事は書けないのである(この号に収録されているものは使用可能by編集部)。

ま、心配ごとはこっちに置いてえ...
...画面を下にスクロールしていくと「御利益 :)」と称して「電話料金、アクセス料金が節約できる」「時間が節約できる」「ゆっくりと読めるので精神衛生上好ましい」などと書いてある(精神衛生のことまで考えていただいて、ドモ、です)。

さらに「利用については回線とサーバーの混み具合を考慮の上ご利用ください。一度に大量のファイルを読むとサーバーにも負荷がかかり、また回線も混雑します。必要ないファイルをダウンロードしないようにしましょう」という気の使いようも好ましい態度かもしれない。

そんなご丁寧なことを言われちゃうと、この下の「利用するには、このライセンス契約書に同意していただく必要があります」なんてとこで、同意しなくっちゃだ。そこで、ここをクリックしてみると、「波乗野郎ダウンロード」というページが……。あれれ、ひょっとしたらこのファイルってモノホンのウェブページをそのまま持ってきちゃったのか? 編集部のCD-ROM担当、某Y氏がシコシコ夜中にダウンロードしているビジュアルが浮かぶではないか.....(合掌)。

いちおう礼儀として文章をちゃんと読み、最後に「同意する」をクリック。と、ト・トツゼンMacPPPが動き出した。げげ~、



「波乗野郎」をインターネットマガジン付録CD-ROMからコピーする。



電話料金、接続料金、時間が節約できると、精神衛生上好ましい?

PPC:Power PC。Power Macintoshなどがこれにあたる。

Readme:アプリケーションの使用方法などが書いてある解説のこと。通常、アプリケーションと同じフォルダーの中に入れられている。



現在のバージョンは漢字Talk7.5以上に対応。



いつのまにかStuffit Expanderがバージョンアップしていた。

InternetConfig：インターネット関連の複数のアプリケーション上で同一の設定ができるインターネット設定ツール。マック版のみ。本誌付録CD-ROMに収録。収録ディレクトリは、

Mac InternetConfig

プリンク：ホームページ上で文字部分が点滅していること。

Stuffit Expander：ドラッグ&ドロップで解凍できるマック用のツール。ウィンドウズ用には、Lhasaという解凍ツールがある。どちらも本誌付録CD-ROMに収録。収録ディレクトリは、

MAC：Mac Stuffit Expander

WIN：WIN LHASA

Bin Hex：バイナリーファイルをテキストに変換するためのマッキントッシュ用プログラム。

な・なんだ？ どうやらリンク先がローカルファイルではなくどっかのURLになっているらしい。

しばらく見てみると、コネクトしたあとで画面にさっき見た「波乗野郎ダウンロード」の画面が……。しかし、こちらは日付が違って「05.29」だ。さらに「現在の最新版」として「Ver.0.90d2 for Macintosh (1996/05/29更新)」の記述がある。わお、もうすでにバージョンアップしててはないか。

早速最新版をダウンロードして……と思ったが、続きの文章をよく読むとMac用の実行環境が「漢字Talk7.5以上」になっていて、「次のリリースから漢字Talk7 (System 7.1～)に対応予定」と書いてある。あれ？ この「次のリリース」ってのは「0.90d2」のこと？ それとも「0.90d2」の次のこと？ ちょっと混乱である。

とりあえず最新版のダウンロードページに行き説明を読むと、どうやらこれが漢字Talk7に対応した「次のバージョン」であることが分かった。う～む、この辺の記述、おっちょこちょいの私のようなユーザーには誤解を生むぞ> 担当者。

🌀 波乗り前のボードのお手入れにちょっと手間取る

続いて説明を読んで行くと、文中に「InternetConfigがあればとても便利です」というところがある。さらに「Macintoshで波乗野郎を利用する際はInternetConfig (フリーウェア)を併用することを強く強くお勧めします」と書いてあり、「強く強く」がプリンクしてるのではないか。

そうまで言われると、知らんぷりもできないので「詳しくはここをご覧ください」をクリック。と「波乗野郎とInternetConfig」(http://www.bug.co.jp/nami-nori/aboutnami-nori.html)が出現。読んでみると、どうやらこれは「Macintosh only」

の機能らしいが、組み合わせて使うと、波乗野郎からネットスケープを立ち上げたり、指定したページを取り込み終わったら、自動的にブラウザが立ち上がって内容を見られたりするという。

日本語版もダウンロードできるようだし、解説図を見る限り、設定も簡単みたいなので、とまかくダウンロードしてみることに……。が、ここで意外なトラブルを体験してしまうことになる。

私の記憶ではこれまでネットスケープを使ってソフトのダウンロードをする場合、ファイルを持ってくると自動的に解凍してくれていたのだが、このファイルの場合はまず「SimpleText」フォーマットのテキストが取り込まれる。では、というわけで「Stuffit Expander」にドラッグ&ドロップして解凍しようとしたところ、こいつがどういうわけかきちんと働かない。途中でファイルが壊れてしまったのかな、と思って再度挑戦したがこれもダメ。そこで「波乗野郎ダウンロード」のページに書いてあった「ダウンロードに失敗したり解凍できなかった場合はこちら」というページに飛んでみると、どうやら「Stuffit Expander4.0.1」が必要らしい。

手元のバージョンをチェックしてみると「3.5.2」だ。いかん。いつの間にやらバージョンが上がってる。インターネットマガジンのCD-ROMをチェックしてみたら、こいつもまだ旧バージョンのまま。仕方がないので、言われるままにAladdin Systemsのページ(http://www.aladdinsys.com/obstufex.htm)から最新版のファイル「stuffit_exp_40_installer.hqx」をダウンロードして「BinHex 4.0」で解凍し、インストール。さらに「InternetConfig日本語版」をもう一度ダウンロード……。(実は「InternetConfig (英語版)」こそがオマケCD-ROMに収録されていたのだが、おマヌケな私はそこまで頭が回らなかったのである。まったくトホホのホ)。

てなことで、余計な手間がかかったが、ともかくこいつのフォルダーに入っている説明ファイル(これもhtmlだ。最近はおっぱらこういうドキュメントが流行なの?)を読む。

それによると、「InternetConfig」はインターネット接続で利用するそれぞれのクライアントソフトが必要とするユーザーの設定をまとめて引き受けてくれるものらしい。これはこれで役立ちモノなんだろうが、今回は主役じゃないので「波乗り野郎」のところで説明されている「InternetConfigの利用方法」に沿って「helper」だけを設定することにした。

⑨ 項目設定でパドリング開始?

意外な伏兵に出会ったおかげで、もう、ひと仕事終わったような気分。英国の探検隊なみにお茶でひと息。体力(知力?)がないとネットサーファーは辛い……。しかし、メインイベントはまだこれからだ。

「波乗り野郎とはの説明のページ(<http://www.bug.co.jp/nami-nori/aboutnami-nori.html>)に「波乗り野郎を詳しく知る」という項目があり、「マニュアル(作成中)近日公開予定 Macintosh用のものがここにあります」とあるので、ここをクリックしてファイルを取り寄せる(ftp://ftp.bug.co.jp/pub/as/nami-nori-doc-mac.sit)。

解凍してみると、結構ドでかい絵が出てくる。でも、なかなかよくできたマニュアルだ。実際に試してみようという人は、取り合えずこのファイルを取り寄せてから始めたほうがいいかもしれない(ただし探検時点で収録されていたものは「0.64d6」用であった)。

まずはニヤけた犬のアイコンをダブルクリックしてプログラムを立ち上げ、「波乗り」メニューか、「PerMan Surfer Data」のボックスの下にある「設定」ボタンを押して「波乗り設定」を行なう【図1】。

設定すべき項目は……

- 1 波乗りするURLの登録
- 2 無視するURLの登録
- 3 波乗りデータ保管フォルダーの場所とProxyの設定
- 4 波乗り開始の方法と終了後の動作の設定

……の4項目。

まずは「追加」ボタン(あるいはウィンドウの上左端の「URL」と書かれたボタン)をクリックするとちょっと大きめのウィンドウが開く【図2】。これはバージョンが新しくなって大きく進化した部分だ。

ここでは大きくわけて3つの設定を行う。1つは、波乗りの対象にするURLの指定。次に、持ってくるファイルのフィルター指定。最後に、もしそのURLにアクセスするのに認証が必要な場合のパスワードなどの設定だ。

最初の「波乗り」の項目では、目的のページの名称とそのURLを入力。さらに「波乗りの対象にする」では、波乗りの対象にするか今回は登録だけしておくかをチェックマークを付けて決定。ここまでは簡単だが、これから先が分かりにくい。

まずは「最大調査リンク数」。ここには数字を入れることになっているが、この数字の意味が、いまいちはっきりしない。あるページにリンクがあった場合、次のリンク先まで行ってファイルを持ってくる。「調査リンク数=1」ということになるのか? う~ん、分らん。こういう場合はデフォルトの「2」でいくことにしよう。

さらに「更新されていないファイルの次を調査しない」というチェックボックスがある。これも意味が分かりにくい。「……ないものを……ない」みたいな、文中に二重の否定があるような文章は作文的には落第だぞ! これをチェックすると、前回アクセスした内容と比較して内容が新しく



ネットサーフィンには体力と知力が必要?



【図1】「波乗り」のメニューから「設定」を選ぶと、この「波乗り設定」の画面になる。



【図2】この画面で指定したURLの設定を行う。

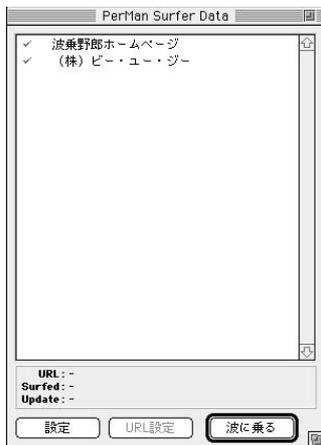
Proxy : クライアントの代理をして相手先のサーバーにアクセスする機能。Proxyを設定しておくと、プロキシサーバーにキャッシュされたデータがクライアントに送られるので、速くアクセスできる。



ずっとサーフィンしているのも疲れるもの……。



【図3】「波乗り開始時間」と「終了後の処理」を設定。



【図4】URLを選んで「波に乗る」をクリックすると、ファイルの収集が開始される。

なっていた場合のみ関連するファイルを取ってくるが、もし更新されていない場合は調査しないということになるのだろうか？ う～ん、それにしても「ファイルの次」って何？ 疑問は残るのである。

次に指定すべきチェックボックスは「目次型ページ」。これは何？ その次の「暗黙ファイル名」というのもよく分からないが、これってURLを指定すると自動的に「index.html」を取り込むってこと？

これ以外の「ファイルフィルタ」の部分では、ページ内にてかきJPEG画像などがあつた場合、時間節約のため持ってこないようにするフィルターだろうし（けど、チェックの場合持ってくるって、フィルターとは逆みたい）、「認証設定」は目的ページでパスワードなどが要求される場合の自動応答だというくらいの想像はつくが、この辺は正式版のマニュアルをちゃんと読む必要がありそうだ。

もとに戻って「波乗り設定」画面の左から2番目、URLに「x」がかかっているボタン。リンク先のデータが大きいので通信料金を節約したいといったときに、アクセスしないURLのリストを作るためのボックスが開く。

左から3番目のボタンは、「波乗り野郎」が取ってきたデータの保管場所とProxyを使っている場合のProxy名とポート番号を入れるウィンドウだ。保管場所指定はMOなどにページを保存したい場合に便利かもしれない。

最後の時計マークのボタンは、プログラム起動後に自動的に「波乗り」を開始するかどうかのチェックボタンと、終了した時の次のアクションの指定だ【図3】。

「波乗りを開始する」をチェックした場合、プルダウンで「すぐに」から「12時間後」までの開始時間指定ができる。

終了後のアクションでは、「なにもしない」「アプリケーション終了」「電源を切る」のほかにも「InternetConfig」が設定され

ていけば「WWW Clientで目次を開く」が選べる。取ってきたファイルをすぐに見たい場合にはこれを選択しておくとも便利である。

以上の設定ができれば、これで大丈夫（なはず）……って、なんか分からない部分があつても、URLとリンク数だけ適当に指定すれば大丈夫みたいだ。

意外に長時間な波乗り野郎のサーフィン

こうして設定した波乗り野郎。早速サーフィンを開始……する前に、ちゃんと手動でPPP接続をしておいたほうがよい。手順としては、通常どおりConfigPPPの「Open」ボタンを押してネットワークに接続してから、波乗り野郎の「波乗り開始」ボタンを押す。これだけで、状態を示す小さなボックスが開き、指定したURLから必要なファイルを持ってきてくれるはずだ【図4】。

ところで、1つのホームページにどれくらいのオブジェクトとリンクがあるかを考えることがあるだろうか？ ホームページを自作した人なら見当がつかかもしれないが、この数は意外に多いのだ。

偉そうなことを言っているが、そんなことはすっかり忘れていた私。スピードの遅いサーバーを指定して、皆様にご迷惑をかけては申し訳ないので、某大手プロバイダのホームページを指定してみた。すると、最初は「o checked 0 got 2 remain」なんてかわいいものだったが、次第に数が増えていき「268 remain」、気がついたらなんと全部で500以上のオブジェクトを持ってきてたのである。最初からオブジェクトの量が多いことを知っていたら、こんな無謀なことはしなかったのに……なんて思ったってもう遅い。

勢いで「気がついたら」と書いてしまったが、正直言って画面に面白い絵も出ない状態ですつと待つというのは、“精神衛生

上” なかなかしんどかった。ま、すべてのファイルを取り込むのは最初のセッションだからで、次の回からは更新されたものだけを持ってくるように設定できるんだから……とも思うが、やっぱり実際には自分が取り込もうとするページがどんなものかを事前にチェックしたり、トライ&エラーで最適な設定を探すしかないのかもしれない。

順調に行けば、あのにやけたワンちゃんが「今日も良い波でした」なんて言いながら、チェックや読み込んだファイル数を報告してくれる（失敗しても「良い波……」と言うところがオマケだが）【図5】。このあとは回線を切り、ブラウザで取ってきたファイルを気の済むまでながめよう。

というわけで簡単に成功したが、最後に気になった点を簡単に報告しておこう。

波乗野郎では、複数の訪問先を登録しておく、ボタン1つで各URLのデータを取ってきてくれる。これは非常に便利な機能なのだが、その中の1つでもアクセスできない（応答が遅くてデータを持ってこれないなど）となると、全プロセスがそこで停止してしまう。実際、これはかなり困る。

次に、セッションが終わっても自動的にモデムが切断されるわけではないので、ずっと数値しか出ないボックスを見続ける必要がある。初回のセッションで取り込むファイル数が多すぎたから、なおさらそう思うのかもしれないが、ついつい途中で止めなくなる。そのために「skip」と「stop」というボタンがあるのだが、これを押してから止まるまでがトロい、トロい。だから口の悪いユーザーに「ばか犬」なんて言われたりするんである（ホントは試用版だから正式に苦情を言ってあげないとね）。

またパソコン通信なんかでは、ターミナルソフトが自動的に電話をかけ、ログインしてデータをダウンロードし終わると、すぐに回線を切断というのが当たり前になっているが、その手の機能がないのはどうも納得いかないな～、という感じである。ひ

よっとしたら、可能なのかもしれないが、うまいやり方は分からなかった。

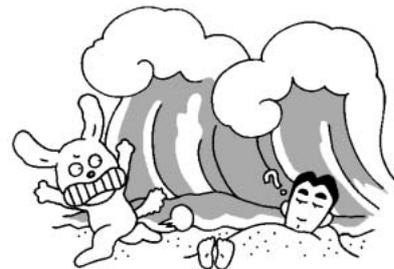
一番問題があるんじゃないかと思ったのは、指定したURLの下のディレクトリーにあるファイルは取りにいくようだが、指定URLのインデックスのポインターが上のディレクトリーを差している場合、ファイルを取りにいけないようなのだ。例えば「立花隆『同時代を撃つ』ホームページ（<http://www.ijinet.or.jp/kodansha/wgendai/index.html>）」の場合、ここのインデックスで出てくる記事は「<http://www.ijinet.or.jp/wgendai/TACHIBANA1-KONONEWS/KONONEWSXXXXXX.html>」みたいな、ちょっと変則的な場所に置かれている。こうした場合はインデックスページを拾ってきてくれるだけで、「どおれ」とのぞいてみると「あれえ、中身がない！」ということになってしまう。

 波乗野郎は慣れてしまうと手放せない？

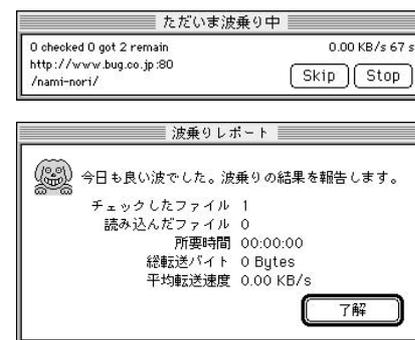
というわけで、試用版ゆえの苦勞はあったが、これらの機能をがちゃんと使えれば、手放せないアイテムになる可能性は十分というのが結論である。

テレホーダイの時間帯、それもたぶん午前2時くらいまでは回線スピードが遅いと思うので、早朝あたりに自動運転で定番ホームページの情報をアップデートする形で使えば、結構便利じゃないだろうか。あるいは、気ままにネットサーフィンしながらそのバックグラウンドで波乗野郎を動かすという手もあるな……。

製品版では、こうした機能がもっと簡単なメニュー形式で実現できているかどうか、近いうちに試してみたいと思う。北海の波乗野郎、潮の加減次第ではなかなか行けるかもしれないということで、今回の探検はおしまいだ。



アクセスできないURLがあると、波乗りがストップしてしまうのは困りもの。



【5】ファイルの収集状況と結果を教えてくれる。

テレホーダイ：一定時間内にあらかじめ決められた相手先にかける電話の料金が月額固定になるNTTの電話サービス。午後11時から午前8時まで同一区域内2か所までの登録ができるテレホーダイ1800と隣接区域2か所までのテレホーダイ3600がある。



[インターネットマガジン バックナンバーアーカイブ] ご利用上の注意

このPDFファイルは、株式会社インプレスR&D(株式会社インプレスから分割)が1994年～2006年まで発行した月刊誌『インターネットマガジン』の誌面をPDF化し、「インターネットマガジン バックナンバーアーカイブ」として以下のウェブサイト「All-in-One INTERNET magazine 2.0」で公開しているものです。

<http://i.impressRD.jp/bn>

このファイルをご利用いただくにあたり、下記の注意事項を必ずお読みください。

- 記載されている内容(技術解説、URL、団体・企業名、商品名、価格、プレゼント募集、アンケートなど)は発行当時のものです。
- 収録されている内容は著作権法上の保護を受けています。著作権はそれぞれの記事の著作者(執筆者、写真の撮影者、イラストの作成者、編集部など)が保持しています。
- 著作者から許諾が得られなかった著作物は収録されていない場合があります。
- このファイルやその内容を改変したり、商用を目的として再利用することはできません。あくまで個人や企業の非商用利用での閲覧、複製、送信に限られます。
- 収録されている内容を何らかの媒体に引用としてご利用する際は、出典として媒体名および月号、該当ページ番号、発行元(株式会社インプレス R&D)、コピーライトなどの情報をご明記ください。
- オリジナルの雑誌の発行時点では、株式会社インプレス R&D(当時は株式会社インプレス)と著作権者は内容が正確なものであるように最大限に努めましたが、すべての情報が完全に正確であることは保証できません。このファイルの内容に起因する直接のおよび間接的な損害に対して、一切の責任を負いません。お客様個人の責任においてご利用ください。

このファイルに関するお問い合わせ先

株式会社**インプレスR&D**

All-in-One INTERNET magazine 編集部

im-info@impress.co.jp