
Subject: Re: Read HTML Page?

Posted by [aqueous0123](#) on Tue, 23 Oct 2001 14:25:47 GMT

[View Forum Message](#) <> [Reply to Message](#)

Holger Nies <nies@nv.et-inf.uni-siegen.de> wrote in message
news:<3BD54399.EAD2FF32@nv.et-inf.uni-siegen.de>...

> Hallo!
> I'm looking for a procedure, I can use to read an Internet page in IDL.
> Is there anybody, who has experience in this area?
> Thank You.
>
> Holger

I know this works under MS Windows, don't know about others.

spawn, "start http://www.usm.edu/index.html"

Subject: Re: Read HTML Page?

Posted by [Logan Lindquist](#) on Tue, 23 Oct 2001 14:29:44 GMT

[View Forum Message](#) <> [Reply to Message](#)

Holger,

You might try developing your GUI in VisualBasic 6, adding IDL functionality as you need it. While there might not be any code written in IDL to parse HTML, there has to be alot of it for other languages. I would also look into C# because like VB6 it offers a faster development process. For a Non-MSFT solution ie, cheaper I would try python. I think you can add COM's (component object model) to that also.

Regards,
Logan

"Holger Nies" <nies@nv.et-inf.uni-siegen.de> wrote in message
news:3BD54399.EAD2FF32@nv.et-inf.uni-siegen.de...

> Hallo!
> I'm looking for a procedure, I can use to read an Internet page in IDL.
> Is there anybody, who has experience in this area?
> Thank You.
>
> Holger
>

Subject: Re: Read HTML Page?

Posted by [Craig Markwardt](#) on Tue, 23 Oct 2001 18:11:42 GMT

[View Forum Message](#) <> [Reply to Message](#)

Holger Nies <nies@nv.et-inf.uni-siegen.de> writes:

> Hallo!
> I'm looking for a procedure, I can use to read an Internet page in IDL.
> Is there anybody, who has experience in this area?
> Thank You.
>
> Holger
>

Hi,

If you have a newer version of IDL, won't the SOCKET procedure work?
This little toy worked for me. Someone at RSI claimed that this
shouldn't work, but it does appear to.

```
socket, unit, 'cow.physics.wisc.edu', 80, /get_lun
printf, unit, 'GET http://cow.physics.wisc.edu/~craigm/idl/idl.html'
text = "
while eof(unit) EQ 0 do begin
    readf, unit, text
    print, text
endwhile
free_lun, unit
```

Craig

--

Craig B. Markwardt, Ph.D. EMAIL: craigmnet@cow.physics.wisc.edu
Astrophysics, IDL, Finance, Derivatives | Remove "net" for better response

Subject: Re: Read HTML Page?

Posted by [Mark Hadfield](#) on Tue, 23 Oct 2001 19:59:02 GMT

[View Forum Message](#) <> [Reply to Message](#)

From: "Holger Nies" <nies@nv.et-inf.uni-siegen.de>

> I'm looking for a procedure, I can use to read an Internet page in IDL.
> Is there anybody, who has experience in this area?

It's clear from the variety of answers you have received already that you

need to tell us exactly what you mean by "read an Internet page".

Mark Hadfield
m.hadfield@niwa.cri.nz <http://katipo.niwa.cri.nz/~hadfield>
National Institute for Water and Atmospheric Research

--

Posted from clam.niwa.cri.nz [202.36.29.1]
via Mailgate.ORG Server - <http://www.Mailgate.ORG>

Subject: Re: Read HTML Page?
Posted by [Andrew Cool](#) on Wed, 24 Oct 2001 00:11:04 GMT
[View Forum Message](#) <> [Reply to Message](#)

Holger Nies wrote:

>
> Hallo!
> I'm looking for a procedure, I can use to read an Internet page in IDL.
> Is there anybody, who has experience in this area?
> Thank You.
>
> Holger

Hallo Holger,

I've emailed you an IDL GUI Web Browser (21kB) that I wrote using IDL's
Socket routine - prompted by Mr Markwardt's post many moons ago.

If you specify just a Web site name, e.g. www.rsinc.com, you'll
get the *basic* HTML for that page, both as text, and displayed
in IE or Netscape. Note, it's simple, basic, and does suck down
all the associated images specified in the HTML.

Alternatively, you can nominate a specific file at a URL, e.g.
a .EXE, or .ZIP, or JPG/GIF/etc, and that file will zoom back,
and if it's a common image format, will hopefully be automagically
displayed.

It even attempts to call your virus scanner software.

Obviously you'll need to edit the code to suit your image display
and virus scanner software.

This package works through the firewall wall here nicely.

Lots of room for fiddling to your heart's content!

Andrew

Andrew D. Cool .->-.
Electromagnetics & Propagation Group `-<-'
Surveillance Systems Division Transmitted on
Defence Science & Technology Organisation 100% recycled
PO Box 1500, Salisbury electrons
South Australia 5108

Phone : 061 8 8259 5740 Fax : 061 8 8259 6673
Email : andrew.cool@dsto.defence.gov.au

Subject: Re: Read HTML Page?

Posted by [Joseph B. Gurman](#) on Wed, 24 Oct 2001 02:45:37 GMT

[View Forum Message](#) <> [Reply to Message](#)

Andrew -

At the risk of being accuysed of a very bad pun, too cool!

Joe Gurman

In article <3BD60718.2AC111FF@dsto.defence.gov.au>, Andrew Cool
<andrew.cool@dsto.defence.gov.au> wrote:

> Holger Nies wrote:

>>

>> Hallo!

>> I'm looking for a procedure, I can use to read an Internet page in IDL.

>> Is there anybody, who has experience in this area?

>> Thank You.

>>

>> Holger

>

> Hallo Holger,

>

>

> I've emailed you an IDL GUI Web Browser (21kB) that I wrote using IDL's

> Socket routine - prompted by Mr Markwardt's post many moons ago.
>
> If you specify just a Web site name, e.g. www.rsinc.com, you'll
> get the *basic* HTML for that page, both as text, and displayed
> in IE or Netscape. Note, it's simple, basic, and does suck down
> all the associated images specified in the HTML.
>
> Alternatively, you can nominate a specific file at a URL, e.g.
> a .EXE, or .ZIP, or JPG/GIF/etc, and that file will zoom back,
> and if it's a common image format, will hopefully be automagically
> displayed.
>
> It even attempts to call your virus scanner software.
>
> Obviously you'll need to edit the code to suit your image display
> and virus scanner software.
>
> This package works through the firewall wall here nicely.
>
> Lots of room for fiddling to your heart's content!
>
>
> Andrew

Subject: Re: Read HTML Page?

Posted by [Andrew Cool](#) on Wed, 24 Oct 2001 03:48:04 GMT

[View Forum Message](#) <> [Reply to Message](#)

"Joseph B. Gurman" wrote:

>
> Andrew -
>
> At the risk of being accuysed of a very bad pun, too cool!
>
> Joe Gurman

>> If you specify just a Web site name, e.g. www.rsinc.com, you'll
>> get the *basic* HTML for that page, both as text, and displayed
>> in IE or Netscape. Note, it's simple, basic, and does suck down
>> all the associated images specified in the HTML.

Alas! Not as "cool" as first described! The word "not" failed to make an appearance between "does" and "suck". So maybe the software does suck, afterall ;-)

The text should read :-

>> If you specify just a Web site name, e.g. www.rsinc.com, you'll
>> get the *basic* HTML for that page, both as text, and displayed
>> in IE or Netscape. Note, it's simple, basic, and does *not* suck down
>> all the associated images specified in the HTML.

Although, with a little fiddling, this too could come to pass...

Andrew

Andrew D. Cool .->-.
Electromagnetics & Propagation Group `-<-'
Surveillance Systems Division Transmitted on
Defence Science & Technology Organisation 100% recycled
PO Box 1500, Salisbury electrons
South Australia 5108

Phone : 061 8 8259 5740 Fax : 061 8 8259 6673
Email : andrew.cool@dsto.defence.gov.au
