Forum:Hack Forums

Thread: Copy a website completely?

Visualpix - *November 3, 2015, 12:15 pm*
Hi, i don't know if this is the right place for the question.

I would like to copy a website completely (or part of it) like uworld or medstudy kaptest.

That is without hacking the website (i don't think it's possible) I have little programming experience but with very little time to do anything else than studying so i don't have time to learn tools like wget, curl (or similar). I need someone to provide/help with me a tool or commands for wget, curl etc.?

uworld.com is what i would like to get most of all. They provide question banks for medical students to prepare for certain examinations. What i've heard is that they restrict possibility to save pages in browser, would there be a way to bypass that. And by copying the site or part of it i mean that i would like the saved offline "site" to be exactly as it would be on the real site - for example in a question with multiple answers, if one "click" on an answer it will save the user input to after the test is done (and also be able to show the correct answers)! Anyone having knowledge about this kind of stuff, please help.

I really don't have much money, but anyone helping with this matter I would do what I can to compensate as much as I can for the work!

Cheers!!

AURUM - *November 3, 2015, 9:44 pm*
I do not think that you can simply copy a website. I have been trying to do this for a very long time without any luck or gain of knowledge. Let me know if you find anything though!

rapta - *November 4, 2015, 12:06 am*
Try httrack httrack .com
it can copy whole website and it's content.

Frosteh - *November 4, 2015, 1:53 am*
You can only copy what you see which is the final rendered HTML that the webserver spits out to you, but that content is constructed from a database of information which you don't have access to, a file store which you may or may not have access to, and a whole bunch of code which does all the logic for how the site works which you don't have access to.

So basically no, you can't copy all of a site including all the functionality, at best you can scrape the site for the style and images/resources it displays, enough to make a trivial but non functioning copy.

Visualpix - *November 4, 2015, 12:11 pm*
Thanks for the replies. I already tried httrack on different sites and I'm guessing (without having any knowledge about these kind of things) that they somewhat have the same configuration as the ones i mentioned before without any success. Frosteh, I get what you're sayin', but hypothetically if anyone with the right skills and the will to copy the site with full functionality it would be possible to do that? And if it so, I'm guessing there is a long learning curve to get to the level of such expertise for a beginner like say 95% percent (maybe more) of the population (including med)!

Cheers!

Baba Yaga - *November 4, 2015, 12:56 pm*
HTTTrack only copies the final render of the HTML page. It does not copy any of the page source codes and any useful information as mentioned in your original post. HTTTrack is mainly used to find and map out a websites directories and such for enumeration purposes.

Nachfolger - *November 4, 2015, 1:00 pm*
1. Download ubuntu
2. open cmd and type "wget "http://domain.com/directory/4?action=AttachFile&do=view&target=file.tgz" "
3. Sit back and enjoy being a hakr

KickAss942 - *November 4, 2015, 8:54 pm*
METHOD TYPE: NO NEED FOR DOWNLOAD ANYTHING AND IT COMPLETES EVERY SOURCE CODE SO EVEN IF YOUR DO NOT HAVE ALL FILES IT PUTS THEM AS URL FILES INTO PROJECT!!!!

Status : Message
WARNING LAZY METHOD DETECTOR - SERVER 3: EASIEST WAY!!!!!!!!!!

Requirments:
Google Chrome & That's IT!!!!

If you're using chrome this is easiest way....

Step 1: Open following urs: chrome://flags/#save-page-as-mhtml
Step 2: Press "Turn on" on the Save Pages as HTML.
Step 3: Restart chrome.
Step 4: Open the site u wanna save.
Step 5: Right click and press > Save As >
Step 6: Change from Filetype = HTML (Website) - Simple File -
to: HTML (Website) Complete (last one).
Step 7: Go to the place u saved it u'll see a folder with all contents used in the page inside the source list and ur main file is the HTML file u saved.

Pretty nice & handy, can rob most websites :D

There are MANY awesome stuff u can change & add using Chrome Flags, but WATCH OUT it puts Chrome into DEVELOPER MODE, meaning you can edit files u're not supposed to edit and settings u're not supposed to see... Which as we all know, could lead to HUGE problems.

Good luck, hope this helped :D

Tyranus. - *November 4, 2015, 8:56 pm*
(Nov 4, 2015 09:54 PM)KickAss942 Wrote: [CITE] -snip-
what the actual fuck was that

KickAss942 - *November 4, 2015, 8:58 pm*
(Nov 4, 2015 09:56 PM)Shadow ?? Wrote: [CITE] what the actual fuck was that
Bruuuuuh, I was editing xD

yloeybht - *November 4, 2015, 9:02 pm*

OP this isn't possible. The website you speak off you need to pay for access. You cannot magically hack the site and steal the content.

Even if you get access to it. They use protective software where they don't give you the correct answers until time is up after taking practice tests.

What you're talking about these companies uses java and other programs to lock up what you're trying to do. They have modules that protects content like that from getting stolen or people to help them cheat on exams. ?

KickAss942 - *November 4, 2015, 9:14 pm*

(Nov 4, 2015 10:02 PM)yloeybht Wrote: [CITE] OP this isn't possible. The website you speak off you need to pay for access. You cannot magically hack the site and steal the content.

Even if you get access to it. They use protective software where they don't give you the correct answers until time is up after taking practice tests.

What you're talking about these companies uses java and other programs to lock up what you're trying to do. They have modules that protects content like that from getting stolen or people to help them cheat on exams. ?

Sorry bruh ur facts are wrong...

Javascript or w.e u use, the files will still show cuz they're NEEDED in the script, and chrome got a timeline in developer tools that tracks every file used & every movement done...

Only possible way u won't get all files are in PHP which u didn't mention so shush :)

Example of PhP page that cannot be saved:
.HTACCESS files in directory = Block.
PHP files with long code lines that has "OnURL = BLANK"
Would show user a blank page but content is full of code that the user cannot see...

Btw ur statements are like a 13yrs skid tht just made me facepalm 25x times cuz I had to write this twice due HF errors....

Anyways Exams? rlly????? SERIOUSLY???? do you EVEN HAVE ANY IDEA why the hell DATABASES WITH PASSWORDS AND USERNAMES AND SERVER IP AND ALL OF THE SECURITY even exists ????

TO SAVE DATA that ONLY you can access as ADMIN, so if there WERE such EXAM ANSWERS to hack, I'd be 2109748901284% sure it won't be INSIDE the script itself, more like GET functions through a DB....

yloeybht - *November 4, 2015, 9:36 pm*

(Nov 4, 2015 10:14 PM)KickAss942 Wrote: [CITE] Sorry bruh ur facts are wrong...

Javascript or w.e u use, the files will still show cuz they're NEEDED in the script, and chrome got a timeline in developer tools that tracks every file used & every movement done...

Only possible way u won't get all files are in PHP which u didn't mention so shush :)

Example of PhP page that cannot be saved:
.HTACCESS files in directory = Block.
PHP files with long code lines that has "OnURL = BLANK"
Would show user a blank page but content is full of code that the user cannot see...

Btw ur statements are like a 13yrs skid tht just made me facepalm 25x times cuz I had to write this

twice due HF errors....
Anyways Exams? rlly????? SERIOUSLY???? do you EVEN HAVE ANY IDEA why the hell DATABASES WITH PASSWORDS AND USERNAMES AND SERVER IP AND ALL OF THE SECURITY even exists ????
TO SAVE DATA that ONLY you can access as ADMIN, so if there WERE such EXAM ANSWERS to hack, I'd be 2109748901284% sure it won't be INSIDE the script itself, more like GET functions through a DB....
Dude, OP wants to copy a website with test exams that are online with answers. These are locked and are server sided. They don't use javascript because you would be able to copy it and cheat.

They use plugins and or java where it locks up PDf'S and other materials.

Many universities have computer exams. The software they use locks up the computer and many different programs when you run their software. The exams run in a locked down system So, no you cannot copy the website 100% including the websites databases and other content without hacking their servers. So, I am right and you're totally wrong with your answer. ?

(Nov 4, 2015 10:14 PM)KickAss942 Wrote: [CITE] Sorry bruh ur facts are wrong...
Javascript or w.e u use, the files will still show cuz they're NEEDED in the script, and chrome got a timeline in developer tools that tracks every file used & every movement done...

Only possible way u won't get all files are in PHP which u didn't mention so shush :)

Example of PhP page that cannot be saved:
.HTACCESS files in directory = Block.
PHP files with long code lines that has "OnURL = BLANK"
Would show user a blank page but content is full of code that the user cannot see...

Btw ur statements are like a 13yrs skid tht just made me facepalm 25x times cuz I had to write this twice due HF errors....
Anyways Exams? rlly????? SERIOUSLY???? do you EVEN HAVE ANY IDEA why the hell DATABASES WITH PASSWORDS AND USERNAMES AND SERVER IP AND ALL OF THE SECURITY even exists ????
TO SAVE DATA that ONLY you can access as ADMIN, so if there WERE such EXAM ANSWERS to hack, I'd be 2109748901284% sure it won't be INSIDE the script itself, more like GET functions through a DB....
You sound really retarded and sound like a skid. You cannot copy a website 100% and just rehost it as your own. You can only copy html and javascript codes. The links can be blocked since it's coming from the server itself. They can block your ass.

OP's site show above is not possible to copy everything all the exams and exercises without paying for membership and even if you paid for it. They have protections on the content itself. The only way is to take a picture of your computer screen. Using your smart phone.

KickAss942 - *November 4, 2015, 9:48 pm*
(Nov 4, 2015 10:36 PM)yloeybht Wrote: [CITE] Dude, OP wants to copy a website with test exams that are online with answers. These are locked and are server sided. They don't use javascript because you would be able to copy it and cheat.

They use plugins and or java where it locks up PDf'S and other materials.

Many universities have computer exams. The software they use locks up the computer and many different programs when you run their software. The exams run in a locked down system So, no you

cannot copy the website 100% including the websites databases and other content without hacking their servers. So, I am right and you're totally wrong with your answer. ?

You sound really retarded and sound like a skid. You cannot copy a website 100% and just rehost it as your own. You can only copy html and javascript codes. The links can be blocked since it's coming from the server itself. They can block your ass.

OP's site show above is not possible to copy everything all the exams and exercises without paying for membership and even if you paid for it. They have protections on the content itself. The only way is to take a picture of your computer screen. Using your smart phone.
Now u're calling me a retard yet ur eyes are still fooling you...

"I would like to copy a website completely (or part of it) like uworld or medstudy kaptest."

He just wanted to "COPY" a website (or PARTS of it) and I gave him a solution, un-happy? try doing something better than nagging over some bs.

Visualpix - *November 7, 2015, 1:07 pm*
(Nov 4, 2015 09:54 PM)KickAss942 Wrote: [CITE] METHOD TYPE: NO NEED FOR DOWNLOAD ANYTHING AND IT COMPLETES EVERY SOURCE CODE SO EVEN IF YOUR DO NOT HAVE ALL FILES IT PUTS THEM AS URL FILES INTO PROJECT!!!!

Status : Message
WARNING LAZY METHOD DETECTOR - SERVER 3: EASIEST WAY!!!!!!!!!!

Requirments:
Google Chrome & That's IT!!!!

If you're using chrome this is easiest way....

Step 1: Open following urs: chrome://flags/#save-page-as-mhtml
Step 2: Press "Turn on" on the Save Pages as HTML.
Step 3: Restart chrome.
Step 4: Open the site u wanna save.
Step 5: Right click and press > Save As >
Step 6: Change from Filetype = HTML (Website) - Simple File -
to: HTML (Website) Complete (last one).
Step 7: Go to the place u saved it u'll see a folder with all contents used in the page inside the source list and ur main file is the HTML file u saved.

Pretty nice & handy, can rob most websites :D

There are MANY awesome stuff u can change & add using Chrome Flags, but WATCH OUT it puts Chrome into DEVELOPER MODE, meaning you can edit files u're not supposed to edit and settings u're not supposed to see... Which as we all know, could lead to HUGE problems.

Good luck, hope this helped :D
Thanks everybody for trying to help. I tested your suggestion Kickass942, thanks alot, it actually gets part of the problem done - saving the webpage content for this kind of website with a very simple trick! Is there a way to save a bigger part of a site or a whole site with this method of saving a webpage? Probably suggestions from everybody else like wget, httrack etc can get the work done as well but I guess I'm to much of an amateur for get it to work.

medstudy free account if someone wants to test.
usermail: oddjob.lebesgue@gmail.com
password: kimyan1

dvrktms - *November 12, 2015, 7:44 am*
Try using pinegrow ?

Here for sample :

h t t p://post img.org/image/4z5bf92jb/full/

zdpooman - *November 12, 2015, 8:02 am*
The problem your going to incur without just asking the owner for a clone environment is the following...

Lets say static HTML is served from one server, while the database is held on another server, and lets go even further down the rabbit hole and say authentication is done through ADFS/SAML on a AD server within their local corporate domain... Your kind of fucked if you think you can clone every appliance, and asset depending on how dynamic or static a website is...

If the content is static, doesn't link out to other assets, and everything resides as HTML with Ajax here and there HTTrack should work fine, just be warned... When you copy a website it doesn't always provide the correct link/file structure and things that you have a "Aha!" moment over could actually just be something missing from a random 403.

?i?d?? - *November 13, 2015, 4:54 pm*
Kali linux distro has a preinstalled hacking software in it. and One of the VERY MANY tools it has preinstalled, can copy a website.