**b.**  **HTTRACK - Offline Browser**

**INTRODUCTION**

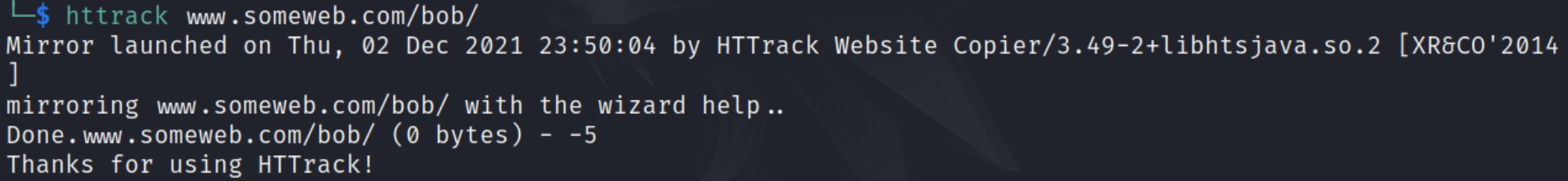
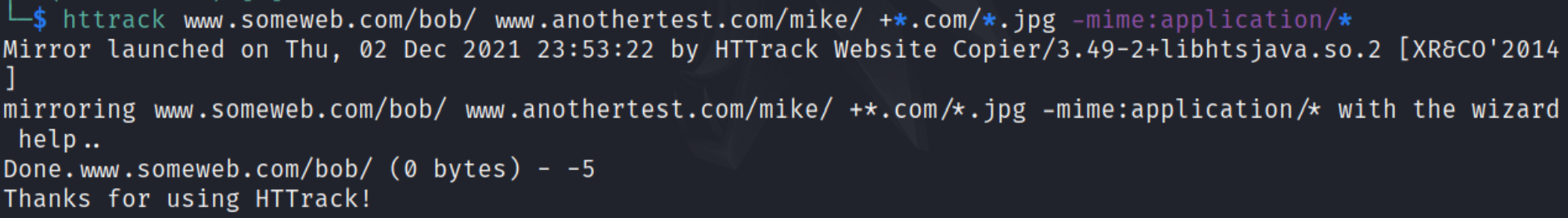
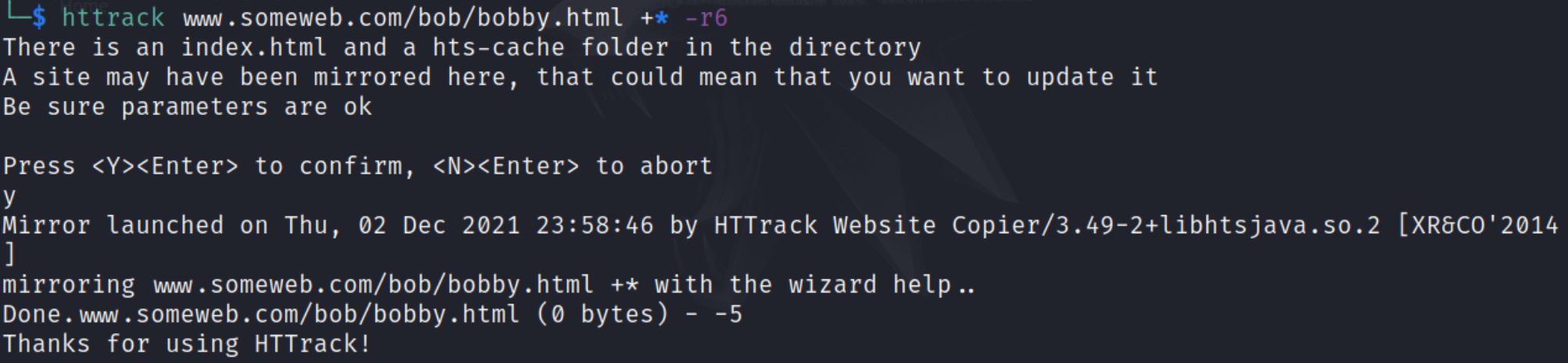
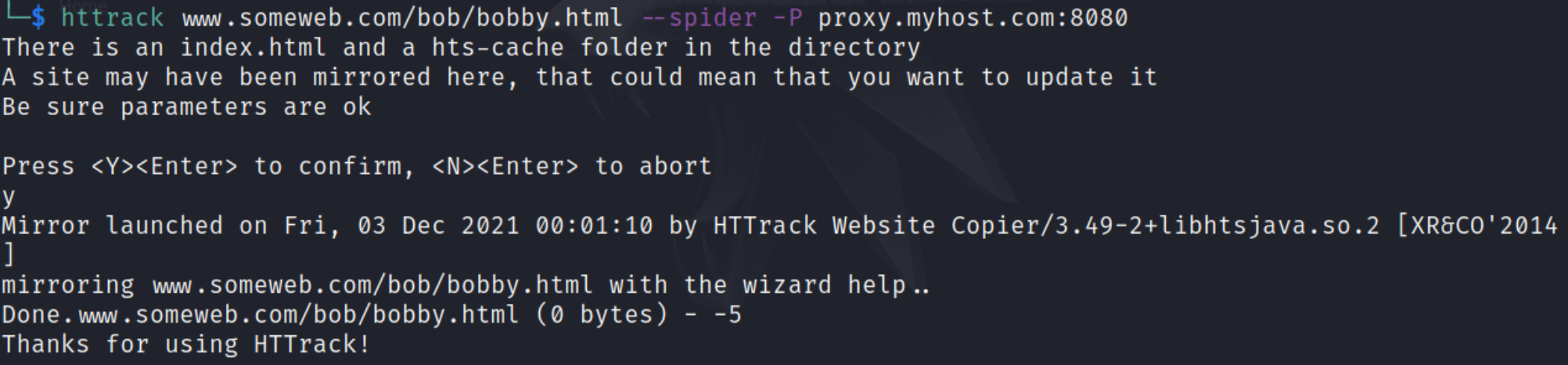
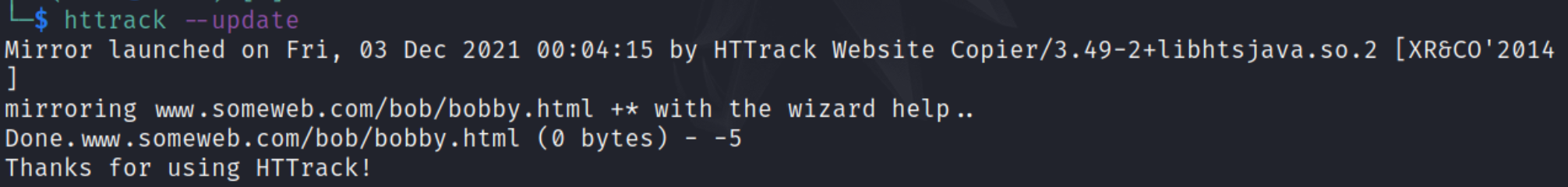
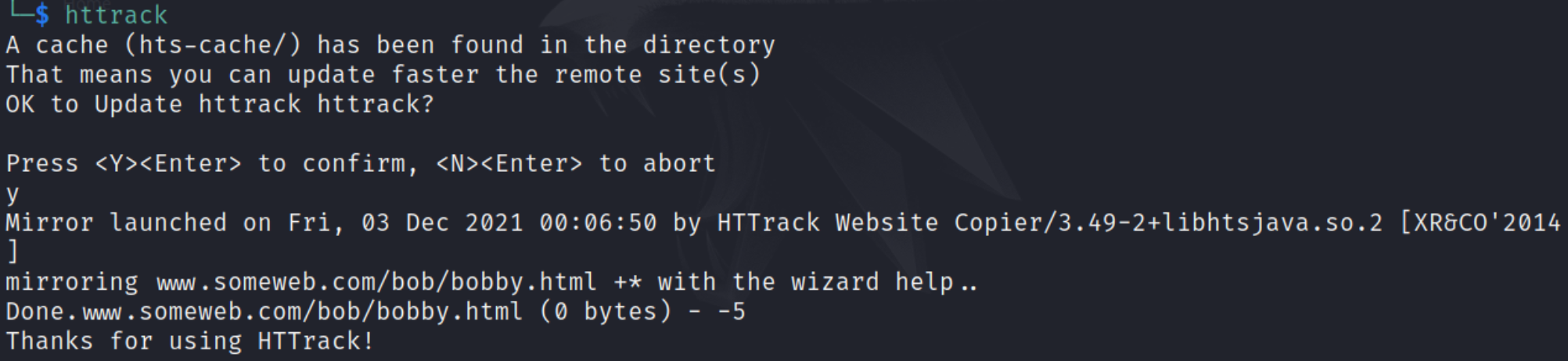
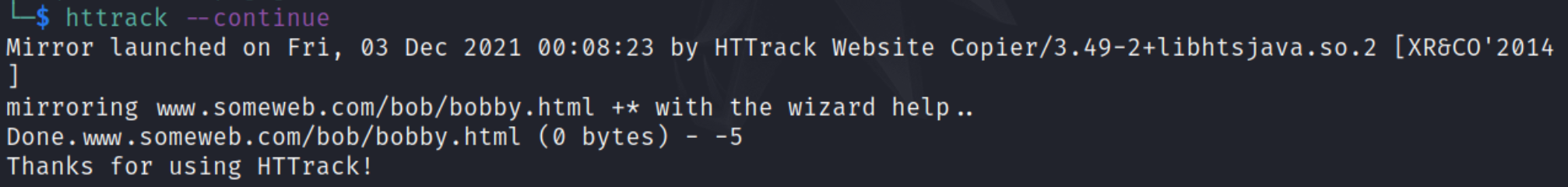
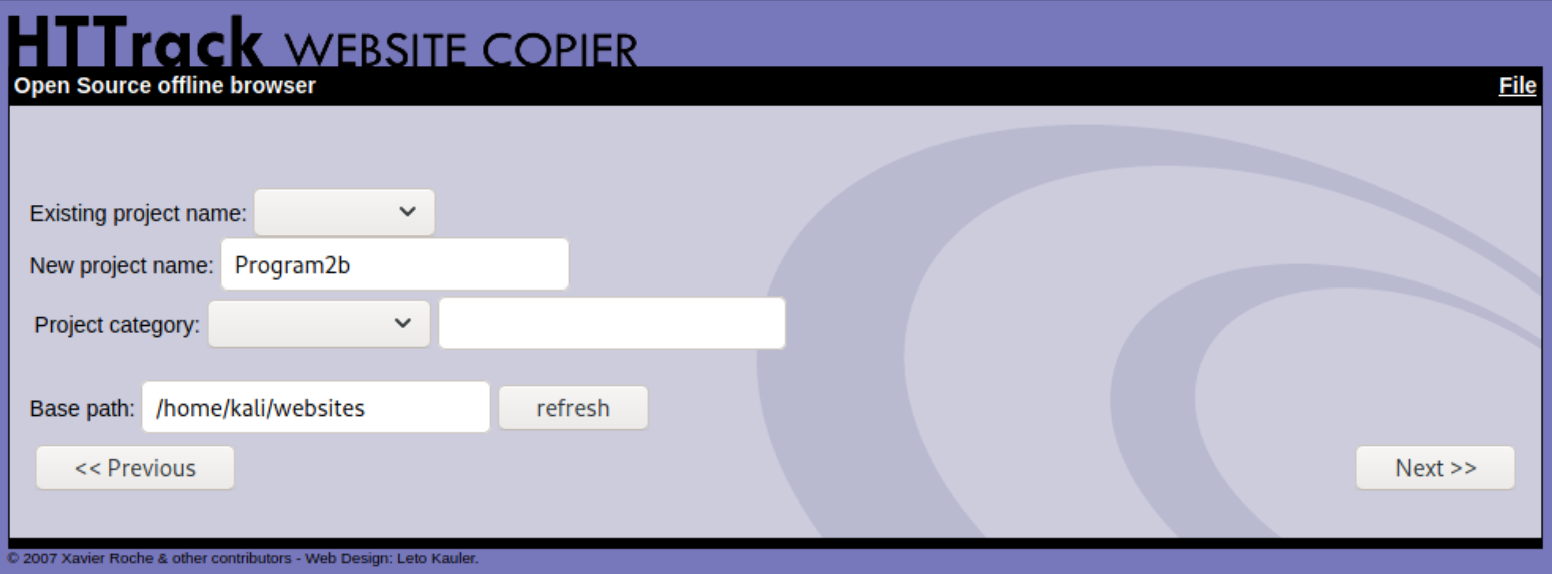
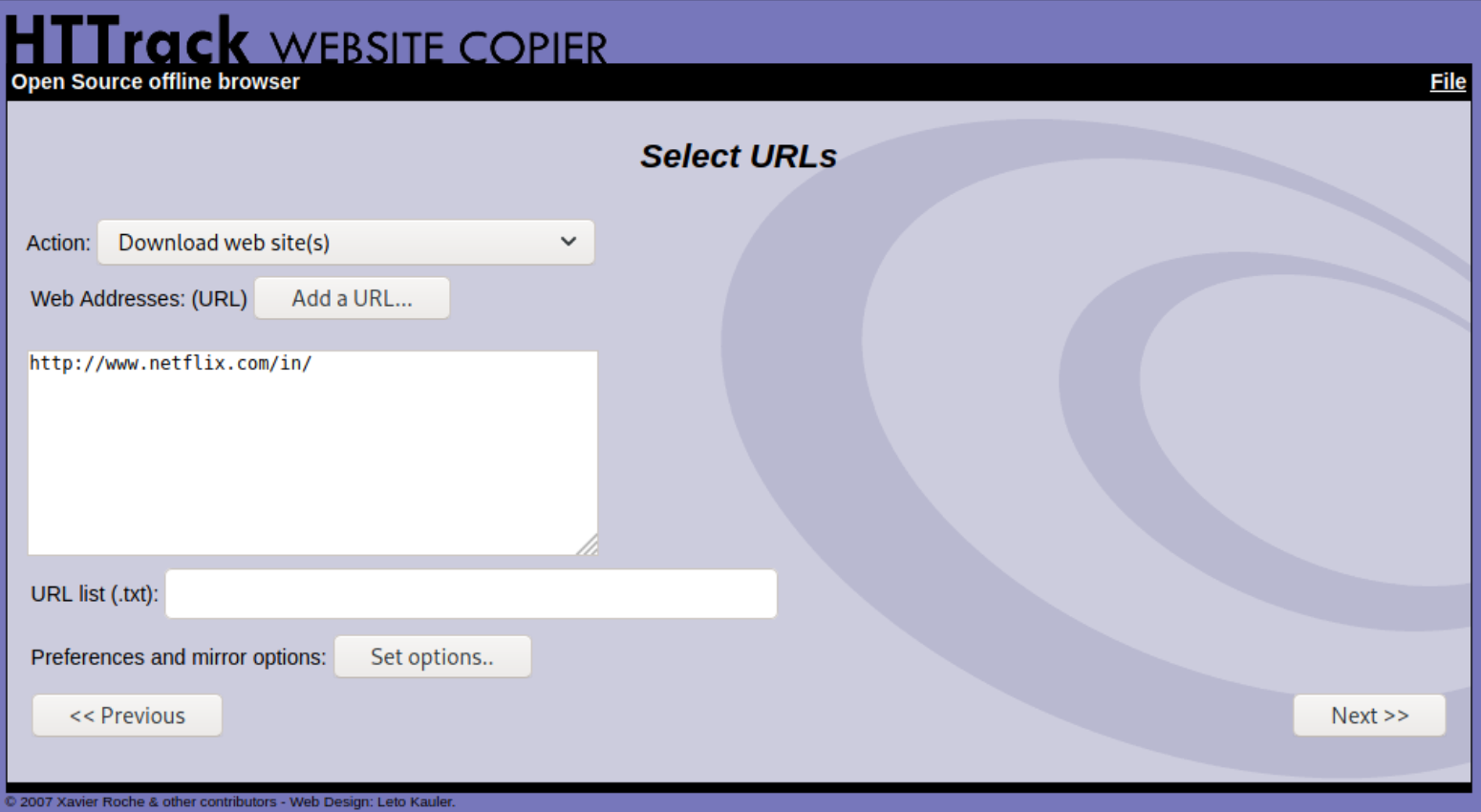
Penetration testing is a simulated cyber attack where professional ethical hackers break into corporate networks to find weaknesses even before attackers do. Httrack is one of the tools that could be used for this purpose.HTTrack is an easy-to-use offline browser utility. It allows you to download a World Wide website from the Internet to a local directory, building recursively all directories, getting html, images, and other files from the server to your computer.   
By default, HTTrack arranges the original site's relative link-structure. Simply open a page of the "mirrored" website in your browser, and you can browse the site from link to link, as if you were viewing it online. HTTrack can also update an existing mirrored site, and resume interrupted downloads. HTTrack is fully configurable, and has an integrated help system.  
HTTrack uses a Web crawler to download a website. Some parts of the website may not be downloaded by default due to the robots exclusion protocol unless disabled during the program. HTTrack can follow links that are generated with basic JavaScript and inside Applets or Flash, but not complex links (generated using functions or expressions) or server-side image maps.

**Objectives -** To clone a World Wide Web ite from the internet to a local computer.

**EXECUTION STEPS**

1. **Installing HTTrack from a package**Command - *sudo apt-get install httrack webhttrack*
2. **Running HTTrack GUI**Command - *webhttrack*
3. **Basic structure**Syntax - *httrack <URLs> [-option] [+<URL\_filter>] [-<URL\_filter>] [+<mime:MIME\_filter>] [-<mime:MIME\_filter>]*

**Examples cases**

1. To mirror a site  
   Syntax - *httrack <site\_URL>  
    httrack* [*www.someweb.com/bob/*](http://www.someweb.com/bob/) **
2. To mirror two sites together (with shared links) and accept any .jpg files on .com sites  
   Syntax - *httrack <site1\_URL> <site2\_URL> +<mime:MIME\_filter> -<mime:MIME\_filter>  
    httrack www.someweb.com/bob/ www.anothertest.com/mike/ +\*.com/\*.jpg -mime:application/\*  
   *
3. To get all files starting from bobby.html, with 6 link-depth, and possibility of going everywhere on the web  
   Command - *httrack www.someweb.com/bob/bobby.html +\* -r6  
   *
4. To run on spider using a proxy  
   Syntax - *httrack <site\_URL> --spider -P <proxy\_link>*   
    *httrack www.someweb.com/bob/bobby.html --spider -P proxy.myhost.com:8080  
   *
5. To update a mirror in the current folder  
   Syntax - *httrack --update  
   *
6. To bring to an interactive mode  
   Syntax - *httrack  
   *
7. To continue a mirror in the current folder  
   Syntax - *httrack --continue  
   ***Using HTTrack GUI**
   1. Syntax - *webhttrack*  
      ****
   2. Create a project  
      
   3. Add the URL  
      
   4. Site mirrored  
      

**CONCLUSION**

1. HTTrack Website Copier is a free open source Web crawler and offline browser software program that allows you to download a copy of a World Wide Web site.
2. It can be a useful tool to make quick backups of a business's websites with its structure intact.

**REFERENCES**

1. HTTrack - <https://en.wikipedia.org/wiki/HTTrack>
2. How to use the HTTrack Website copier - <https://smallbusiness.chron.com/use-httrack-website-copier-52794.html>
3. HTTrack: Penetration Testing Tools - <https://en.kali.tools/?p=443>