Title : - Accessing the sensitive data of a website using robots.txt

Description : - Robots.txt is a standard used by websites to communicate with web crawlers and other web robots. The standard specifies how to inform the web robot about which areas of the website should not be processed or scanned.
For more information on robots.txt file visit this [link]

Summary : - By using the robots.txt file we can access the part of the website where the creator doesn't want us to look and can also retrieve sensitive data from the website.

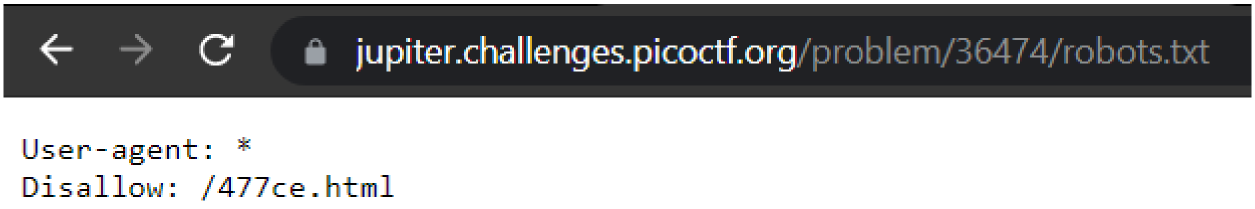
Steps to Reproduce : -
Step 1 - Open the link https://jupiter.challenges.picoctf.org/problem/36474/
Step 2 - In the address bar after the link type robots.txt and now we can see the User-agent and Disallow.
Step 3 - Remove the robots.txt after the address and copy paste the Disallow (/477ce.html) after the address and press enter
Step 4 - and now we got the flag, paste the flag in the picoCTF Submit box.

Payload : - robots.txt
/477ce.html

Impact : - By using the robots.txt file the sensitive information present on the website can be easily accessed by the hackers.

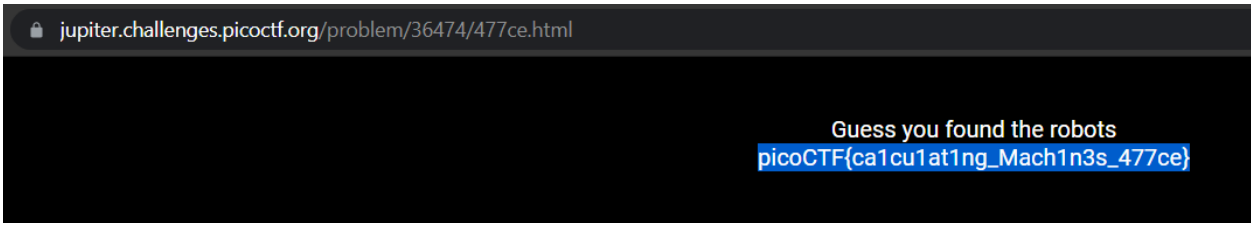
Mitigation :- Make your site more secure and scan the request coming from the web crawlers (robots.txt)

Proof of concept : -



jupiter.challenges.picoctf.org/problem/36474/robots.txt

```
User-agent: *
Disallow: /477ce.html
```



jupiter.challenges.picoctf.org/problem/36474/477ce.html

Guess you found the robots
picoCTF{ca1cu1at1ng_Mach1n3s_477ce}