Practical non-image based CAPTCHA approaches?

Asked 16 years, 4 months ago   Modified 12 years, 5 months ago

Viewed 85k times

316

votes

🔒 **Locked**. This question and its answers are [locked](#) because the question is off-topic but has historical significance. It is not currently accepting new answers or interactions.

It looks like we'll be adding [CAPTCHA](#) support to Stack Overflow. This is necessary to prevent bots, spammers, and other malicious scripted activity. We only want human beings to post or edit things here!

We'll be using a JavaScript (jQuery) CAPTCHA as a first line of defense:

[http://docs.jquery.com/Tutorials:Safer_Contact_Forms_Without_CAPTCHAs](#)

The advantage of this approach is that, **for most people, the CAPTCHA won't ever be visible!**

However, for people with JavaScript disabled, we still need a fallback and this is where it gets tricky.

I have written a [traditional CAPTCHA control for ASP.NET](#) which we can re-use.

However, I'd prefer to go with something textual to avoid the overhead of creating all these images on the server with each request.

I've seen things like..

- ASCII text captcha: `\/\/(_)\/\/`

- math puzzles: what is 7 minus 3 times 2?

- trivia questions: what tastes better, a toad or a popsicle?

Maybe I'm just tilting at windmills here, but I'd like to have a less resource intensive, non-image based `<noscript>` compatible CAPTCHA if possible.

Ideas?

security   language-agnostic   captcha

Share   edited Apr 1, 2012 at 9:23

community wiki

16 There is no need to actually create an image on the server. You just need to handle the request. For example <img src="generateImage.aspx?guid=blah"> – Brian R. Bondy Oct 19, 2008 at 4:44

58 Trivia questions are prone to cultural bias (think of a french guy answering your question...). Furthermore, they can tackle users whose English isn't native. Also, they can easily be broken using brute force (you only have ~2^#_OfQuestions options). – Adam Matan Jan 26, 2009 at 9:29

72 Also, what on earth is a popsicle? – Fraser Mar 14, 2009 at 2:06 ✎

57 According to Wolfram Alpha, "what is 7 minus 3 times 2" is 1. I thought it was 8. I think you just invented the anti-captcha. – Mike Robinson Jan 14, 2010 at 22:55 ✎

50 @Mike Robinson: I think programmers should know about operator precedence in NORMAL day use =) – Gnark Feb 10, 2010 at 10:04

Comments disabled on deleted / locked posts / reviews |

103 Answers

Sorted by: Highest score (default) ⇕

Prev | 1 | 2 | 3 | 4 | Next

2 votes

🔖

🕘

I am sure most of the pages build with the controls (buttons, links, etc.) which supports mouseovers.

- Instead of showing images and ask the user to type the content, ask the user to move the mouse over to

any control (pick the control in random order (any button or link.))

- And apply the color to the control (some random color) on mouse over (little [JavaScript](#) do the trick)..

- then let the user to enter the color what he/she has seen on mouse over.

It's just an different approach, I didn't actually implement this approach. But this is possible.

Share

edited Mar 29, 2010 at 12:59

community wiki
2 revs, 2 users 74%
RameshVel

2

votes

Just be careful about cultural bias in any question based CAPTCHA.

[Bias in Intelligence Testing](#)

Share

edited Mar 29, 2010 at 13:04

community wiki
2 revs, 2 users 80%
Cameron MacFarland

2

votes

Use a simple text CAPTCHA and then ask the users to enter the answer backwards or only the first letter, or the last, or another random thing.

Another idea is to make a ASCII image, like this (from Portal game end sequence):

```
                                  ., - - -.
                                ,/XM#MMMX;,
                              -%###########M%,
                             -@######%   $###@=
             ., --,         -H#######$    $###M:
          ,;$M###MMX;        .;##########$;HM###X=
        ,/@##########H=         ;###############+
       -+#############M/,         %##############+
       %M###############=          /#############:
       H################          .M#############;.
       @###############M          ,@###########M:.
       X################,          -$=X########@:
       /@#################%-          +######$-
       .;#################X         .X#####+,
        .;H###############/         -X####+.
          ,;X#############,           .MM/
            ,:+$H@M#######M#$-         .$$=
                ., -=;+$@###X:        ;/=.
                    .,/X$;       .::,
                       .,         ..
```

And give the user some options like: IS A, LIE, BROKEN HEART, CAKE.

Share   edited Apr 19, 2010 at 16:17

                                              community wiki

2

votes

My solution was to put the form on a separate page and pass a timestamp to it. On that page I only display the form if the timestamp is valid (not too fast, not too old). I found that bots would always hit the submission page directly and only humans would navigate there correctly.

Won't work if you have the form on the content page itself like you do now, but you could show/hide the link to the special submission page based on NoScript. A minor inconvienience for such a small percentage of users.

Share

edited Apr 19, 2010 at 16:20

community wiki
3 revs, 2 users 63%
dsims

2

votes

The fix-the-syntax-error CAPTCHA:

```
echo "Hello, world!;
for (int $i = 0; $i < 10; $i ++ {
  echo $i /*
}
```

The parens and quotes are randomly removed.

Bots can automatically check syntax errors, but they don't know how to fix them!

Share

answered Nov 21, 2010 at 22:51

community wiki
Ming-Tang

1

vote

@pc1oad1etter I also noticed that after doing my post. However, it's just an idea and not the actual implementation. Varying the font or using different colours instead of bold/italics would easily address usability issues.

Share

answered Aug 12, 2008 at 5:47

community wiki
Jarod Elliott

1

vote

@rob

> What about a honeypot captcha? Wow, so simple! Looks good! Although they have highlighted the accessibility issue.. Do you think that this would be a problem at SO? I personally find it hard to imagine developers/programmers that have difficulty

> reading the screen to the point where they
> need a screen reader?

There are developers who are not just legally blind, but 100% blind. Walking cane and helper dog. I hope the site will support them in a reasonable fashion.

However, with the honeypot captcha, you can put a hidden div as well that tells them to leave the field blank. And you can also put it in the error message if they do fill it in, so I'm not sure how much of an issue accessibility really is here. It's definitely not great, but it could be worse.

Share

answered Aug 12, 2008 at 6:29

community wiki
Derek Park

1

vote

How about showing nine random geometric shapes, and asking the user to select the two squares, or two circles or something.. should be pretty easy to write, and easy to use as well..

There's nothing worse than having text you cannot read properly...

answered Aug 14, 2008 at 13:25

community wiki
Thomi

1

vote

Have you looked at [Waegis](#)?

"Waegis is an online web service that exposes an open API (Application Programming Interface). It gets incoming data through its API methods and applies a quick check and identifies spam and legitimate content on time. It then returns a result to client to specify if the content is spam or not."

answered Aug 15, 2008 at 16:46

community wiki
Jon Galloway

1

vote

I think they are working on throttling. It would make more sense just to disable CAPTCHA for users with 500+ rep and reset the rep for attackers.

answered Aug 15, 2008 at 19:09

community wiki

1

vote

I recently (can't remember where) saw a system that showed a bunch of pictures. Each of the pictures had a character assigned to it. The user was then asked to type in the characters for some pictures that showed examples of some category (cars, computers, buildings, flowers and so on). The pictures and characters changed each time as well as the categories to build the CAPTCHA string.

The only problem is the higher bandwidth associated with this approach and you need a lot of pictures that are classified in categories. There is no need to waste much resources generating the pictures.

Share

edited Aug 15, 2008 at 20:27

community wiki
2 revs
HS.

1

vote

My suggestion would be an ASCII captcha it does not use an image, and it's programmer/geeky. Here is a PHP implementation http://thephppro.com/products/captcha/ this one is a paid. There is a free, also PHP implementation, however I could not find an example -> http://www.phpclasses.org/browse/package/4544.html

I know these are in PHP but I'm sure you smart guys building SO can 'port' it to your favorite language.

answered Aug 22, 2008 at 9:09

community wiki
UnkwnTech

1

vote

Answering the original question:

- ASCII is bad : I had to squint to find "WOW". Is this even correct? It could be "VVOVV" or whatever;

- Very simple arithmetic is good. Blind people will be able to answer. (But as Jarod said, beware of operator precedence.) I gather someone could write a parser, but it makes the spamming more costly.

- Trivia is OK, but you'll have to write each of them :-(

I've seen pictures of animals [what is it?]. Votes for comics use a picture of a character with their name written somewhere in the image [type in name]. Impossible to parse, not ok for blind people.

You could have an audio fallback reading alphanumerics (the same letters and numbers you have in the captcha).

Final line of defense: make spam easy to report (one click) and easy to delete (one recap screen to check it's a spam

account, with the last ten messages displayed, one click to delete account). This is still time-expensive, though.

Share

1
vote

Our form spam has been drastically cut after implementing the honeypot captcha method as mentioned previously. I believe we haven't received any since implementing it.

Share

1
vote

Perhaps the community can come up with some good text-based CAPTCHAs?

We can then come up with a good list based on those with the most votes.

Share

1
vote

[Mollom](#) is another [askimet](#) type service which may be of interest. From the guys who wrote drupal / run acquia.

Share

answered Sep 2, 2008 at 20:58

1
vote

I think we must assume that this site will be subject to targeted attacks on a regular basis, not just generic drifting bots. If it becomes the first hit for programmers' searches, it will draw a **lot** of fire.

To me, that means that any CAPTCHA system **cannot pull from a repeating list of questions**, which a human can manually feed into a bot, in addition to being unguessable by bots.

Share

answered Sep 9, 2008 at 18:20

1

vote

One way I know of to weed out bots is to store a key in the user's cookie and if the key or cookie doesn't existing assume they're a bot and ignore them or fall back in image CAPTCHA. It's also a really good way of preventing a bunch of sessions/tracking being created for bots that can add a lot of noise to your DB or overhead to your system performance.

Share

answered Oct 7, 2008 at 15:55

community wiki
iamgoat

1

vote

One thing that is baffling is how Google, apparently the company with the most CS PHDs in the world can have their Captcha broken, and seem to do nothing about it.

Share

answered Oct 8, 2008 at 10:22

community wiki
Chris S

1

vote

You don't only want humans posting. You want humans that can discuss programming topics. So you should have a trivia captcha with things like:

What does the following C function declaration mean: `char *(*(**foo [][8])())[];` ?

=)

Share

It means someone's got too much time on their hands! (I wouldn't be able to answer it - I don't do C) – Andrew Grimm Dec 21, 2010 at 7:33

1

vote

Which color is the fifth word of this sentence? red?, blue, green?

(color words adequately)

Share

1

vote

I think a custom made CAPTCHA is your best bet. This way it requires a specifically targeted bot/script to crack it. This

effort factor should reduce the number of attempts. Humans are lazy afterall

Share

answered Nov 6, 2008 at 2:43

community wiki
Andrew Harry

1

vote

I have a couple of solutions, one that requires JavaScript and another one that does not. Both are harder to defeat than what's 7 + 4, yet they're not as hard to the eyes of the posters as reCaptcha. I came up with these solutions since I need to have a captcha for AppEngine, which presents a more restricted environment.

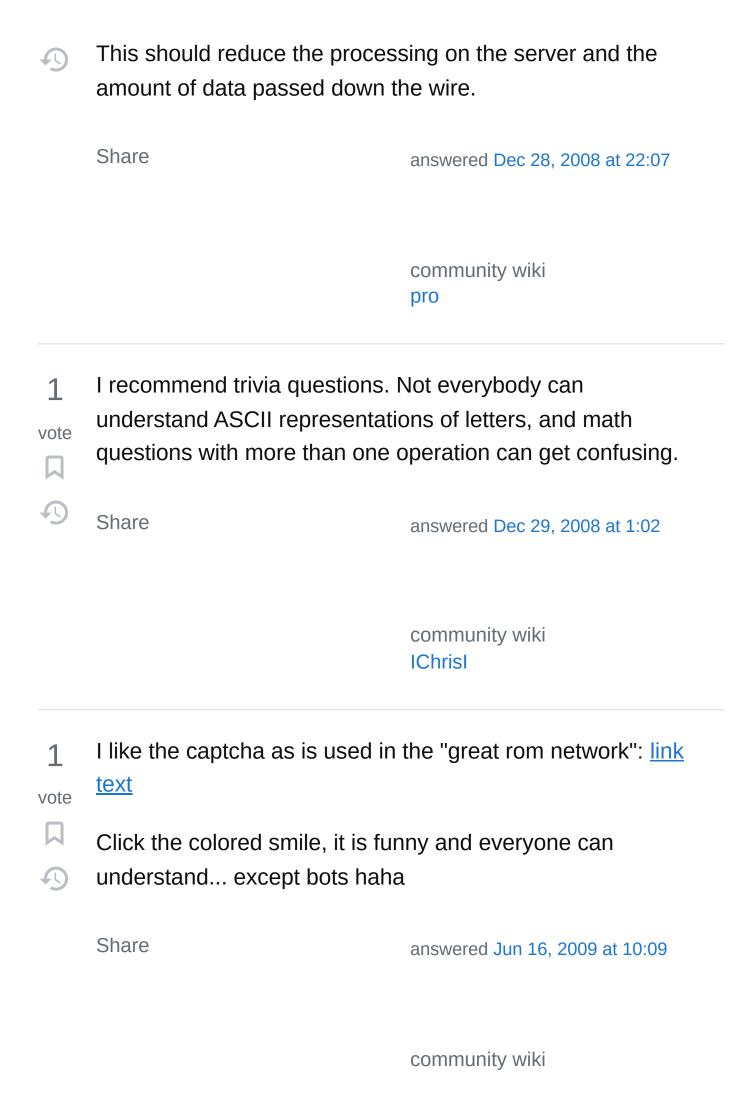Anyway here's the link to the demo: http://kevin-le.appspot.com/extra/lab/captcha/

Share

answered Nov 11, 2008 at 8:42

community wiki
Kevin Le - Khnle

1

vote

The image could be created on the client side from vector based information passed from the server.

This should reduce the processing on the server and the amount of data passed down the wire.

Share

answered Dec 28, 2008 at 22:07

community wiki
pro

1

vote

I recommend trivia questions. Not everybody can understand ASCII representations of letters, and math questions with more than one operation can get confusing.

Share

answered Dec 29, 2008 at 1:02

community wiki
IChrisI

1

vote

I like the captcha as is used in the "great rom network": link text

Click the colored smile, it is funny and everyone can understand... except bots haha

Share

answered Jun 16, 2009 at 10:09

community wiki

1

vote

Just to throw it out there. I have a simple math problem on one of my contact forms that simply asks

what is [number 1-12] + [number 1-12]

I probably get probably 5-6 a month of spam but I'm not getting that much traffic.

Share

answered Jul 9, 2009 at 3:15

community wiki
[corymathews](#)

1

vote

I really like the method of captcha used on this site: http://www.thatwebguyblog.com/post/the_forgotten_timesaver_photoshop_droplets#commenting_as

Share

answered Jul 10, 2009 at 5:24

community wiki
[jsnfwlr](#)

1

vote

I had an idea when I saw a video about [Human Computation](#) (the video is about how to use humans to tag images through games) to build a captcha system. One could use

such a system to tag images (probably for some other purpose) and then use statistics about the tags to choose images suitable for captcha usage.

Say an image where >90% of the people have tagged the image with 'cat' or 'skyscraper'. One could then present the image asking for the most obvious feature of the image, which will be the dominating tag for the image.

This is probably out of scope for SO, but someone might find it an interesting idea :)

Share

answered Aug 24, 2009 at 20:41

community wiki
Runeborg

1

vote

Here's my captcha effort:

```
The security number is a spam prevention measure and
is located in the box
of numbers below. Find it in the 3rd row from the
bottom, 3rd column from
the left.

208868391    241766216    283005655    316184658
208868387    241766212

241766163    283005601    316184603    208868331
241766155    283005593

241766122    283005559    316184560    208868287
241766110    283005547
```

```
316184539    208868265    241766087    283005523
316184523    208868249

208868199    241766020    283005455    316184454
208868179    241766000

316184377    208868101    241765921    283005355
316184353    208868077
```

Of course the numbers are random as is the choice of row and collumn and the choice of left/right top/bottom. One person who left a comment told me the 'security question sucks dick btw':

http://jwm-art.net/dark.php?p=louisa_skit

to see in action click 'add comment'.

Share

answered Oct 22, 2009 at 21:20

community wiki
James Morris

@chris, no I'm just not a trends whore. – James Morris Apr 10, 2011 at 16:39