No CAPTCHA reCAPTCHA

EECS 3482 - Introduction to Computer Security

Tom Bassakyros
Dennis Nada
Adebayo Abdulahi Olalekan

CAPTCHA

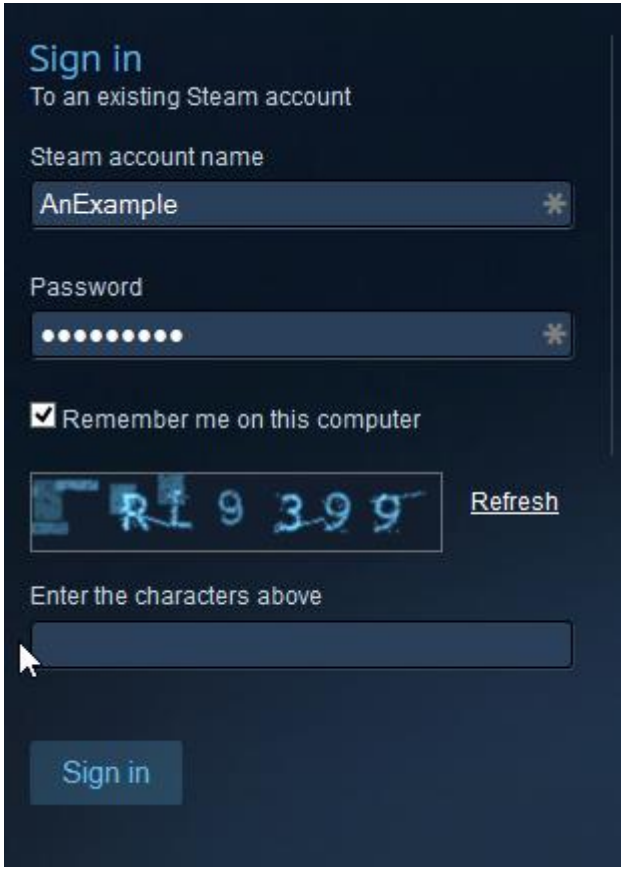
- **C**ompletely **A**utomated **P**ublic **T**uring test to tell **C**omputers and **H**umans **A**part

History

- Process was first patented in 1998 by Mark D. Lillibridge, Martin Abadi, Krishna Bharat, and Andrei Z. Broder at AltaVista
- Came up with the idea by looking for ways to make it harder for Optical Character Recognition (OCR) to read the text in their images
- Some factors OCR used to read text where similar typefaces and plain backgrounds. They designed their image riddle by doing the complete opposite.

Applications

- Preventing comment spam on sites
- Protecting Website Registration
- Preventing Dictionary Attacks

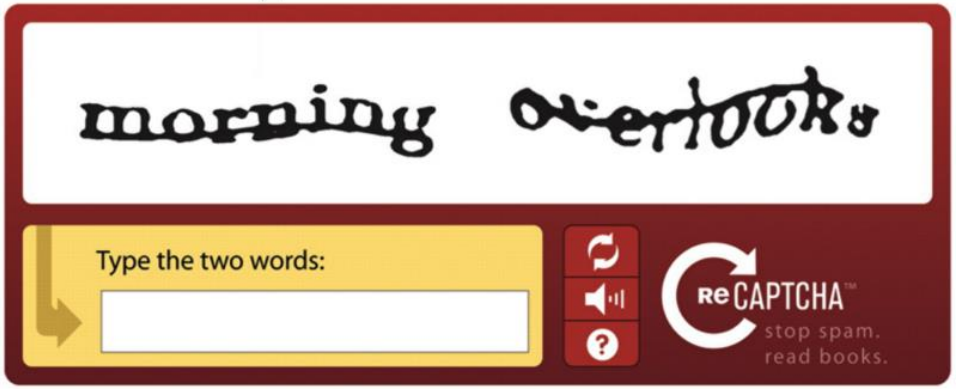


Problems with CAPTCHA

- Hard to read, especially with visually impaired users
- Although audio CAPTCHAs are available, some find it almost as hard to decipher
- As technology advances, newer OCR programs are available to read CAPTCHAs.
  - Example: Z-Gimpy has a success rate of 92%

reCAPTCHA

- Developed by Luis von Ahn, later acquired by Google in 2009
- Unlike the CAPTCHA system, it presents two words instead of one. One of the word is the CAPTCHA word and the other is a word that needs to be digitized.
- In 2012, reCAPTCHA started using photographs of house number to help aid Google's Street View



How it works

- A dialogue is displayed with distorted text images and prompts the user to write the distorted text for “human validation” which is to differentiate bots from humans.
- Scanned text is canned by two OCR programs, if a word is deciphered differently by both OCR programs or that is not in the English dictionary is marked as "suspicious" and converted into a CAPTCHA.
- If the human types the control word correctly, then the response to the questionable word is accepted as probably valid.
- The identification performed by each OCR program is given a value of 0.5 points, and each interpretation by a human is given a full point. Once a given identification hits 2.5 points, the word is considered valid.

Advantages

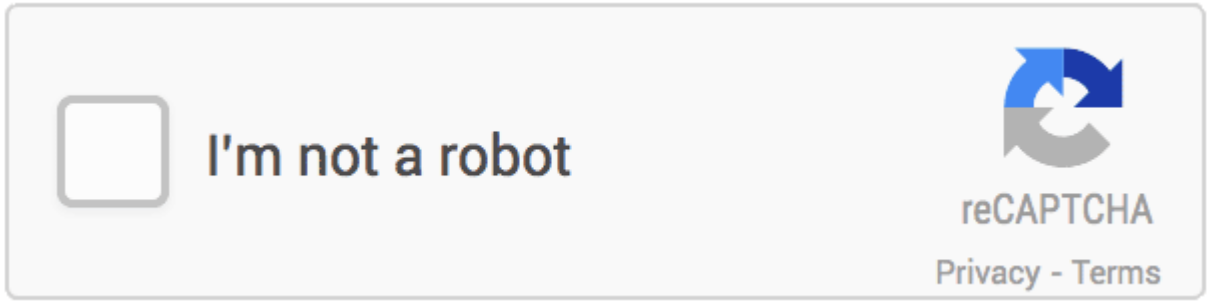
- More advanced security than the old system as reCAPTCHA is constantly evolving with Google's involvement.
- reCAPTCHA offers more than just spam protection. Every time our CAPTCHAS are solved, that human effort helps digitize text, annotate images and build machine-learning datasets.

Disadvantages

- By adding two words that are hard to read, it becomes less accessible. This makes it hard for visually-impaired people to pass the reCAPTCHA.
- More time is needed to decipher the reCAPTCHA
- Technical difficulties with certain web browsers.

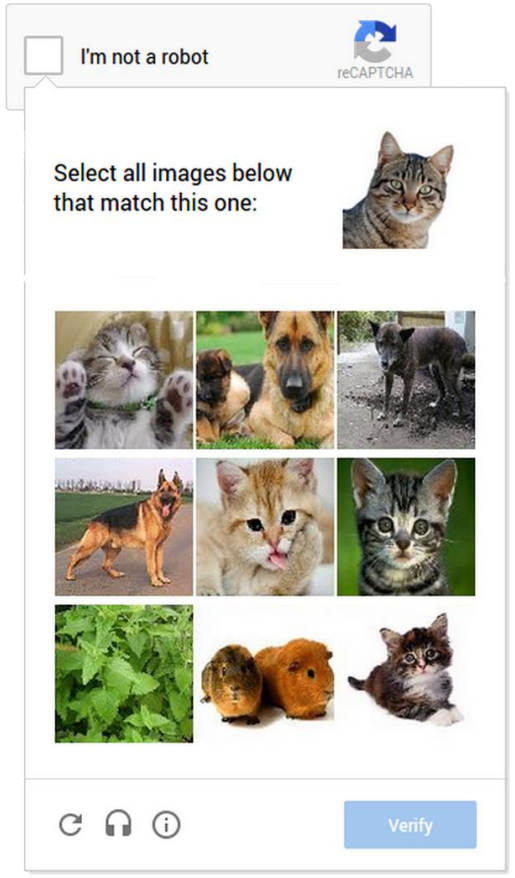
No CAPTCHA reCAPTCHA

- Introduced in December 2014, the new system would only require user to click the checkbox to verify the user
- Uses a sophisticated backend analysis called “Advanced Risk Analysis” to consider the user’s engagement **before, during and after** clicking the checkbox.
- Some known factors include user’s IP address, browser cookies, and even mouse movement.



Mobile No CAPTCHA reCAPTCHA

- No CAPTCHA reCAPTCHA does not use a checkbox for mobile browsers, instead uses a “image quiz”
- Ex. It will display a picture of a cat and you would need to select the images that are also cats.

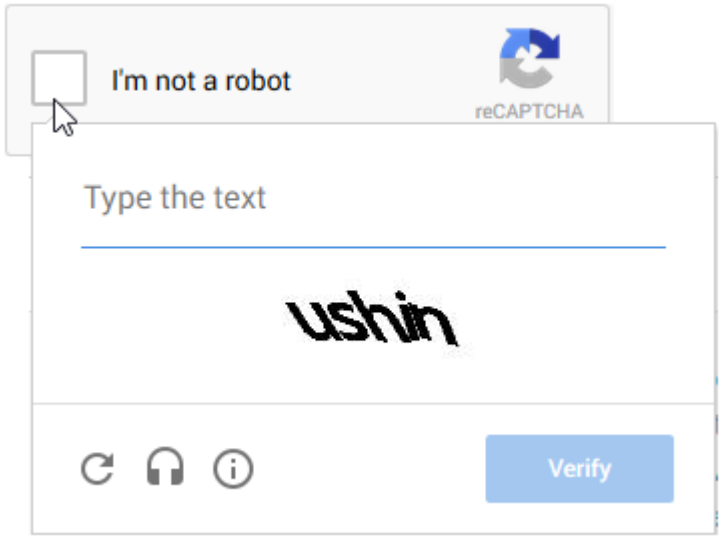


Advantages

- Better accessibility, less “squiggly characters” makes it better for the visually-impaired.
- Images reCAPTCHA has better passing rate than the old system

Disadvantages

- If cookies are not kept by the browser (such as using Incognito or Private browsing), the old reCAPTCHA system is used.



Question 1

- **Name at least one example of an application for CAPTCHAs.**
    - Preventing comment spam on sites
    - Protecting Website Registration
    - Preventing Dictionary Attacks.

Question 2

- **Which of the following is not prompted by reCAPTCHA?**

A. Random Numbers
B. Random Alphabetical Letters
C. Pictures of House Numbers
D. Dictionary Words

Question 3

- **Which of the following is not a main factor during the Advanced Risk Analysis of the No CAPTCHA ReCAPTCHA?**

A. Browser Cookies
B. Mouse Cursor Movements
C. External Devices
D. IP Address

References

1. http://googleonlinesecurity.blogspot.ca/2014/12/are-you-robot-introducing-no-captcha.html
2. http://www.wired.com/2014/12/google-one-click-recaptcha
3. http://venturebeat.com/2014/12/24/googles-no-captcha-recaptchas-may-not-be-as-bot-proof-as-we-thought/
4. http://simplyaccessible.com/article/googles-no-captcha/
5. https://www.google.com/recaptcha
6. http://lanena19-informationtechnology.blogspot.ca/2010/10/captcha-advantages-and-disadvantages.html
7. https://www.cylab.cmu.edu/partners/success-stories/recaptcha.html
8. http://www.captcha.net/
9. http://blog.drwile.com/?p=10743
10. http://www.techspot.com/news/53495-disabled-australian-starts-petition-to-kill-captcha.html
11. http://download.springer.com/static/pdf/990/bok%253A978-3-540-39200-2.pdf?auth66=1424400793_262d1abaa7ec00dd5b0add020704a0ac&ext=.pdf
12. http://www.cs.sfu.ca/~mori/research/gimpy/

Thank you!