

# [Everything You Ever Wanted To Know About CAPTCHAs But Were Afraid To Ask \[Technology Explained\]](#)

[James Bruce](#) June 3, 2011

Love them or hate them – CAPTCHAs have become ubiquitous on the Internet. What is CAPTCHA anyway, and where did it come from? Responsible for eye-strain the world over, the humble CAPTCHA has been the centre of much attention as the single most effective weapon in the fight against web spam. But are they effective? Are there any other kinds of CAPTCHA other than the basic “tilt your head, squint and read me”?

## **You Shall Not Pass (The Turing Test)!**

Captcha’s were invented by a team of Carnegie Mellon professors and put into first use around the year 2000 by AltaVista and Yahoo, in an attempt to prevent automated chat bots and URL submissions. **It is in fact an acronym for Completely Automated Public Turing test to tell Computer and Humans Apart.**

But what is a Turing test? Named after British professor Alan Turing, the Turing test is the standard test of an *Artificial Intelligence* based machine, whereby if a machine can pass the test, it is considered to exhibit *intelligent behaviour*. Essentially the test involves conversing with a number of judges through a text interface – if the judges can’t tell they are chatting to a computer, it passes the test.



The CAPTCHA therefore, is an automated Turing test. There are a number of different ways of doing this, but the most common one that we seem to have settled on is to present the user with a scrambled form of text, assuming (often incorrectly) that any normal human will be able to decipher the text. The CAPTCHA has evolved over time, but has ultimately been defeated as we’ll find out later.

## Verify Your Registration

• Enter the code shown:

[More info](#)

This helps prevent automated registrations.



## Text-Based CAPTCHAs & The Re-CAPTCHA Project

The reCAPTCHA project, now owned by Google, decided that instead of inately deciphering cryptic text for no real good, it presented a fantastic opportunity to correct the shortcomings of computer-based **Optical Character Recognition**. For older books especially, computers find it very hard to recognise the words, whereas a human finds the tasks trivial. Combine the task of digitising old books with spam prevention, and you're onto an absolute winner.



However, if the computer had trouble recognising the word in the first place, how can it tell if what you wrote in is nonsense? Simple – *present the user with TWO words* – one of which is known. The system assumes that if the user correctly types the known word, then the chances are that the unrecognisable word is also correct. Another ingenious idea is to [combine the CAPTCHA with some form of advertising](#).

## Math Problem

### Qualifying question

Just to prove you are a human, please answer the following math challenge.

Q: Calculate:

$$\frac{\partial}{\partial x} \left[ 4 \cdot \sin \left( 7 \cdot x - \frac{\pi}{2} \right) \right] \Big|_{x=0}$$

A:

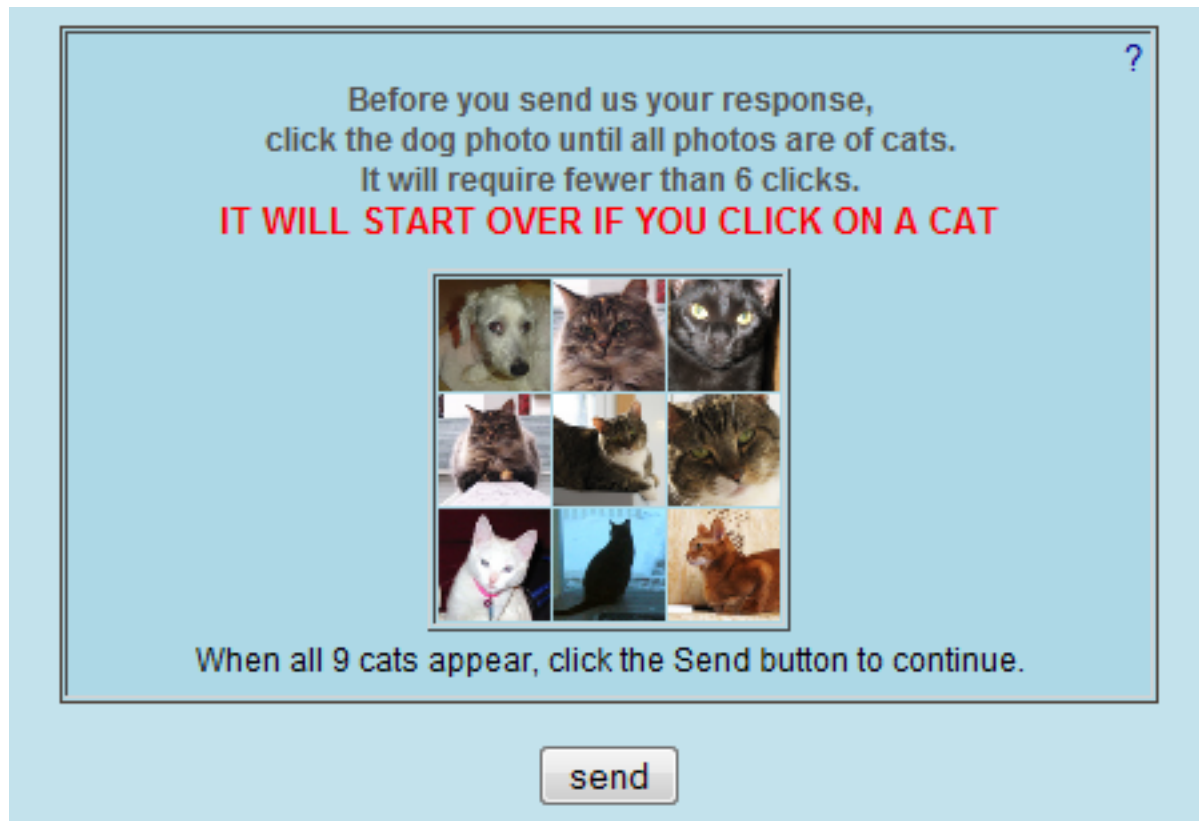
mandatory

Note: If you do not know the answer to this question, reload the page and you'll get another question.

OK, the picture is a joke, but essentially the user is presented with a basic math problem. We use a similar system on the [Answers site](#) right now. It needn't be difficult, just some basic addition.

## Image-Based CAPTCHAs

As difficult as some of the ReCAPTCHA codes can be for you and I sometimes, software has already been developed which can break the code with about a 30% success rate – which for a spam campaign with millions of tries is quite an acceptable rate. **Images on the other hand are extremely difficult to process for computers semantically. Think about a simple cat picture – programming a computer to recognise a human face is hard enough, but to distinguish a cat from all the other animals and objects in the world is pretty much impossible at this point in time.**



## Logic-Based

These rely on logical and semantic intelligence about the world, or just basic common human sense. Some examples might be:

- Identify the food in this list: asphalt, bacon, cloud, dagger.
- Identify the weapon in this list: asphalt, bacon, cloud, dagger.
- How many doors are on a four-door car?
- What is the third word in this sentence?
- What's left if you remove the B from ABC?

A great plugin to integrate these kind of tests into your WordPress comment system is [WP-Gatekeeper](#), by the way.

## De-CAPTCHA Services

**The sad fact is that while CAPTCHAs are a necessary evil, they are easily overcome by spammers nowadays.** While some spammers have indeed developed [sophisticated software](#) that can mimic the human eye and brain to decode like a human does, the truth is far more simpler and more horrific. Why develop expensive software when you can pay someone pennies to do the CAPTCHA for you? **The current cheapest going rate is \$1.39 for 1000 CAPTCHAs,** with a 98% accuracy rate, and services such as [Death By Captcha](#) have developed elaborate APIs for developers to use. The only person being slowed down by CAPTCHAs nowadays, is you!

## The Future Of The Captcha

Like everything else in life, CAPTCHAs are not impenetrable to hacking or spamming. As new and more ingenious tests are devised, ever more sophisticated ways of breaking them will be developed – and the solution of paying someone else to do them for you can never be defeated. Even so, it's our responsibility as web developers and admins to keep spammers away from our sites without degrading user experience.

Image Credit : [xkcd](#)