



Prove You Are Human

As we have seen here at LadyBug in the Writer's Room, automatic spammers can wreak havoc with sites that allow posts from any source. They can post the most idiotic and obscene things, requiring constant vigilance. After a while the effort becomes exceedingly frustrating and then totally overwhelming.

One possible solution you probably have seen, a box with distorted letters that you have to type in before you can have access, is quite prevalent at places like Yahoo Groups and Gmail. The idea is that a human can read those letters, while a computer can't. A **CAPTCHA** (Completely Automated Public Turing test to tell Computers and Humans Apart) attempts to make sure that the computer desiring access to a site is operated by a human being. This can foil various automated attacks on sites like blogs and internet polls.

You may remember the Turing test from the early years of computing... a person sits in front of a terminal asking questions and tries to decide whether the answer she has received is from a computer or a fellow human being. Artificial intelligence people felt successful if the computer's answer sounded like that of a human being to the questioner.

The CAPTCHA process is similar, but with a twist. Now the questioner who declares the answerer to be a human being is a computer. The computer says, in effect, "Hey, you are able to read something a computer couldn't read, so you must be human! Come on in!"

Putting all the deciphered passwords typed in at the many web sites that require them is quite a total human effort. Reportedly 150,000 hours a day are used figuring them out. So the folks at Carnegie Mellon had a great idea...in reCAPTCHA they use this decoding effort to help supplement OCR efforts while digitizing books. Optical Character Recognition (OCR) is about 84 % accurate. So they have two OCR programs digitize the same material. If they agree on a word, fine. If they don't, that word gets put into a CAPTCHA together with a known word and presented to users several times. If enough of them agree, that word is used. The final product has accuracy of 99%, that of excellent human transcribers. Right now we users are working on, among other things, old editions of the New York Times.

Isn't it great to be human?

Live example of a [reCAPTCHA](#)

Overview article:

Science, Volume 321, 12 September 2008, pp 1465-1468