



MAKERERE

UNIVERSITY

COLLEGE OF COMPUTING AND INFORMATION SCIENCES
DEPARTMENT OF COMPUTER SCIENCE
COURSEWORK: RESEARCH METHODOLOGY(BIT 2207)
LECTURER: ERNERST MWEBAZE

Fighting spam with reCAPTCHAs

ODEKE MOSES
216016894

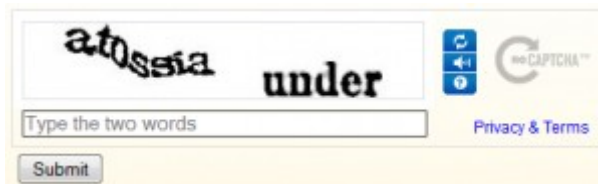
Contents

1	INTRODUCTION	2
2	CLAIM	2

1 INTRODUCTION

Human Interactive Proofs (HIPs) focus on automation tests that virtually all humans can pass but current computer programs fail. Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHA) was an acronym that was coined in 2000. It is a type of challenge-response test that only a human completes successfully. In the simplest form of a CAPTCHA, an image consisting of digits and letters sufficiently distorted is presented and the user is required to input the characters that are displayed. Other forms of CAPTCHAs are based on text-graphics, audio, hand-writing and puzzles. CAPTCHAs have been widely used as a security measure to restrict access from Robots or Bots. CAPTCHAs are based on Artificial Intelligence (AI) problems that cannot be solved by current computer programs or Bots but are easily solvable by humans. A client who provides a correct response to a challenge is presumed to be a human; otherwise a Bot.[1]

Luis von Ahn is a computer scientist who was instrumental in developing the CAPTCHA back in the late 90's and early 2000's. According to an article the Canadian magazine The Walrus, when CAPTCHAs started to become popular, Luis von Ahn "realized that he had unwittingly created a system that was frittering away, in ten-second increments, millions of hours of a most precious resource: human brain cycles."



An example of a reCAPTCHA CAPTCHA.

In order to try and ensure that this time was not wasted, von Ahn set about developing a way to better utilise this time; it was at this point that reCAPTCHA was born.

2 CLAIM

reCAPTCHA is a free service that protects your website from spam and abuse. reCAPTCHA uses an advanced risk analysis engine and adaptive CAPTCHAs to keep automated software from engaging in abusive activities on your site. It does this while letting your valid users pass through with ease.

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. This in turn helps preserve books, improve maps, and solve hard AI problems.

Since computers have trouble reading squiggly words like these, CAPTCHAs are designed to allow humans in but prevent malicious programs from scalping tickets or obtain millions of email accounts for spamming. But there's a twist, the words in many of the CAPTCHAs provided by reCAPTCHA come from scanned archival newspapers and old books. Computers find it hard to recognize these words because the ink and paper have degraded over time, but by typing them in as a CAPTCHA, crowds teach computers to read the scanned text[2]. In this way, reCAPTCHA's unique technology improves the process that converts scanned images into plain text, known as Optical Character Recognition (OCR). This technology also powers large scale text scanning projects like Google Books and Google News Archive Search. Having the text version of documents is important because plain text can be searched, easily rendered on mobile devices and displayed to visually impaired users.

reCAPTCHA has completed digitizing the archives of The New York Times and books from Google Books, as of 2011. The archive can be searched from the New York Times Article Archive, where more than 13 million articles in total have been archived, dating from 1851 to the present day.[3] Through mass collaboration, reCAPTCHA was helping to digitize books that are too illegible to be scanned by computers, as well as translate books to different languages, as of 2015.

References

- [1] "A FRAMEWORK FOR DEVANAGARI SCRIPT-BASED CAPTCHA", *Sushma Yalamanchili and Kameswara Rao*, No. 4, August 2011.
- [2] "Teaching computers to read: Google acquires reCAPTCHA" , *Google Official Blog*, September 16 2009.
- [3] " ReCAPTCHA", *<https://en.wikipedia.org/wiki/ReCAPTCHA>*, [Accessed: 08- Mar- 2018].