

Программы против спама помогут "отцифровать" книги

Компьютерные программы, защищающие от спама, сделали возможным с большей точностью перевести в электронную форму старинные книги и другие письменные источники.

Получить доступ ко многим сайтам можно лишь зарегистрировавшись на них, предварительно напечатав в специальном окошке набор символов, представленных в виде причудливой картинке. Владельцы интернет-ресурсов идут на это для противодействия натиску спам-сообщений.

Сейчас многие сайты вместо невнятной абракадабры используют для этих целей тексты из старинных письменных источников, которые были отсканированы и но не распознаны.

Куски текста, которые программы распознавания текста не осилили, как раз и предлагается расшифровать интернет-пользователям.

С учетом того, что около 100 млн. пользователей в день доказывают таким образом, что они не роботы, получается довольно внушительная армия потенциальных дешифровальщиков.

Объем нечитаемых слов в некоторых старинных документах может достигать до 20% из-за пожелтевшей бумаги или выцветших чернил.

Предельная точность

Специалисты из американского университета Карнеги-Меллон разработали систему под названием Recaptcha, которая собирает фрагменты сканированного текста, не поддавшиеся компьютерной программе, и затем рассылает их сайтам, которые подписаны на эту услугу.

Вместе с буквами, которые компьютер не узнал после сканирования, высылаются и слово, уже внесенное в систему. Правильно распознав второе слово, пользователь получает доступ к сайту, а расшифровка первого слова дополняет базу данных, после того, как будет получено несколько подтверждающих ответов.

В материале, опубликованном в журнале Science, говорится, что такой подход позволяет добиться почти стопроцентной точности, что превосходит требования архивариусов.

В настоящее время около 40 тыс. сайтов предлагают своим пользователям решить такие простенькие задачки, и получают в день около 4 млн. ответов.

В прошлом году таким образом удалось расшифровать 440 млн. слов и перевести в цифровую форму весь архив газеты New York Times с 1908 года.

[👍 Порекомендовать текст](#)

Поделиться в соцсетях:

Также читайте эксклюзивную информацию в соцсетях:

- [Телеграм](#)

- [ВКонтакте](#)

Связаться с редакцией Бабра:

newsbabr@gmail.com

НАПИСАТЬ ГЛАВРЕДУ:

Телеграм: @babr24_link_bot
Эл.почта: newsbabr@gmail.com

ЗАКАЗАТЬ РАССЛЕДОВАНИЕ:

эл.почта: bratska.net.net@gmail.com

КОНТАКТЫ

Бурятия и Монголия: Станислав Цырь
Телеграм: @bur24_link_bot
эл.почта: bur.babr@gmail.com

Иркутск: Анастасия Суворова
Телеграм: @irk24_link_bot
эл.почта: irkbabr24@gmail.com

Красноярск: Ирина Манская
Телеграм: @kras24_link_bot
эл.почта: krasyar.babr@gmail.com

Новосибирск: Алина Обская
Телеграм: @nsk24_link_bot
эл.почта: nsk.babr@gmail.com

Томск: Николай Ушайкин
Телеграм: @tomsk24_link_bot
эл.почта: tomsk.babr@gmail.com

[Прислать свою новость](#)

ЗАКАЗАТЬ РАЗМЕЩЕНИЕ:

Рекламная группа "Экватор"
Телеграм: @babrobot_bot
эл.почта: equatoria@gmail.com

СТРАТЕГИЧЕСКОЕ СОТРУДНИЧЕСТВО:

эл.почта: babrmarket@gmail.com

[Подробнее о размещении](#)

[Отказ от ответственности](#)

[Правила перепечаток](#)

[Соглашение о франчайзинге](#)

[Что такое Бабр24](#)

[Вакансии](#)

[Статистика сайта](#)

[Архив](#)

[Календарь](#)

[Зеркала сайта](#)