

3. Google's reCAPTCHA service often requires website users to respond to a reCAPTCHA prompt by typing two words that are displayed as two distorted images. One of those words is a "known" word, which the website user must enter correctly to access the website as a security measure. That is, because Google already knows what word is being displayed in the first distorted image, if the user enters the word correctly, Google knows the user is likely to be a human, and thus permits the users to continue using the website. If the user does not enter the word correctly, there is a greater chance that the user is not a human and may be accessing the website for illegitimate reasons; Google thus denies access if the user enters the word incorrectly (although often providing a second chance for the user if the first attempt results in an incorrect answer).

4. The other of the two words, however, serves no security purpose. The second word is an image with text that Google is attempting to transcribe. The sole purpose of the second word is to require the user to read and transcribe the word for Google's commercial use and benefit, with no corresponding benefit to the user.

5. Google transcribes images in exchange for payments from third parties, such as the New York Times (which uses reCAPTCHA to transcribe archived images of New York Times articles), and Google transcribes images for its own commercial use, such as for its Google Books and Google Street View services, as explained more fully below.

6. Although Google derives substantial profit from the transcriptions of images generated by users of websites that employ reCAPTCHA, Google does not disclose to website users that Google is profiting from their time and effort. Nor does Google compensate users for such time and effort. In essence, Google operates a highly profitable transcription business built upon free labor, which it deceptively and unfairly obtains from unwitting website users, unjustly enriching itself at the expense of website users responding to reCAPTCHA prompts. Google's conduct is also a violation M.G.L. c. 93A, § 2 ("Chapter 93A").

7. Plaintiff seeks, on behalf of herself and other Massachusetts residents who have responded to a reCAPTCHA prompt, to recover damages, including statutory and multiple damages, interest, attorney's fees, costs, and any and all other relief permitted by law.

Parties

8. Plaintiff Gabriela Rojas-Lozano is a resident of Williamstown, Massachusetts. In August 2014, while living in Massachusetts, Ms. Rojas-Lozano signed up for a new Gmail account through Google's Gmail service, <http://www.gmail.com>. As a condition to creating a new Gmail account, Google required Ms. Rojas-Lozano to respond to a reCAPTCHA prompt by typing two words that were displayed to Ms. Rojas-Lozano as two distorted images. The second of those two words served no security purpose; the sole purpose of the second word was to require Ms. Rojas-Lozano to read and transcribe the word for Google's use and benefit. Ms. Rojas-Lozano did not receive any compensation for transcribing the second word.

9. Defendant Google is a Delaware corporation that maintains its headquarters in Mountain View, California.

Jurisdiction and Venue

10. This Court has jurisdiction pursuant to 28 U.S.C. §§ 1332(d) and 1453, because (1) this action is a "class action," which seeks certification of proposed classes of individuals; (2) the putative Classes consist of at least hundreds of thousands of class members; (3) the citizenship of at least one class member is different from Google's citizenship; and (4) the aggregate amount in controversy by the claims of Plaintiff and the putative Classes exceeds \$5,000,000, exclusive of interest and costs, given the number of class members in this case and the potential for actual and statutory damages under Chapter 93A.

11. This Court also has personal jurisdiction over Google because Google does business in Massachusetts, and the actions that are the subject of this complaint took place in substantial part in Massachusetts.

12. Venue is proper in this jurisdiction pursuant to 28 U.S.C. § 1391 because Google is subject to personal jurisdiction in this District, and a substantial portion of the conduct complained of herein occurred in this District.

Factual Allegations

CAPTCHA Security Prompts

13. A “CAPTCHA” is a website security feature that requires a website user, before accessing a website or certain features of a website, to manually type letters and/or numbers that are displayed to the user in a distorted image. (CAPTCHA is an acronym for “Completely Automated Public Turing test to tell Computers and Humans Apart.”)

14. The stated purpose of CAPTCHAs is to ensure that website users are human beings rather than an automated program that may not be accessing a website for legitimate reasons. That is, because automated programs are generally unable to decipher distorted images of texts (absent deployment of sophisticated optical character recognition technology (“OCR”)), but human users can typically decipher distorted images, a CAPTCHA prompt can distinguish real human beings from automated programs, preventing the latter from accessing all or parts of websites that use the CAPTCHA service.

15. There are a number of companies offering CAPTCHA solutions to website owners that provide security to website owners wishing to exclude automated programs from some or all of a particular website.

Google's reCAPTCHA Platform

16. Google's CAPTCHA platform, reCAPTCHA, is the dominant CAPTCHA provider and has been widely used on popular websites including Facebook, Twitter, TicketMaster, CNN, and Craigslist, as well as Google's own websites including Gmail.

17. Most CAPTCHAs require that a user type one word after viewing a distorted image of that word in order to proceed past the CAPTCHA prompt. An example of a one-word CAPTCHA, generated by a platform other than Google's reCAPTCHA, is provided below:



18. Google's reCAPTCHA is unique in that it often requires users to type two separate words displayed as distorted images in order to proceed past the reCAPTCHA prompt. An example of such a two-word Google reCAPTCHA is provided below:



19. Where Google requires a user to type two separate words displayed as distorted images, one of the two words serves no security purpose. Rather, Google requires the website user

to decipher and enter the second word to further Google's own commercial purposes, thereby enabling Google to profit from the user's deciphering the second distorted image.

20. Google has used its reCAPTCHA platform to transcribe scanned images of books and newspapers, and also to decipher addresses found in images captured as part of Google's Street View project.

21. In the course of Google's digitization of books, newspapers, and other images, there are words or number series that Google's OCR technology cannot decipher accurately. For this reason, OCR technology alone is insufficient to create an accurate transcription of scanned texts. Human transcription remains necessary for complete and accurate transcriptions of such texts.

22. Google uses humans to decipher words or number series through reCAPTCHA prompts. By adding a second distorted image in the reCAPTCHA prompt that multiple users are required to decipher and enter, Google is able to determine the correct transcription of the word or number series without having to pay any human to determine the correct transcription. This improves the quality of Google's digitization of newspapers, books, and other texts.

23. For example, in the image between paragraphs 18 and 19 above, one of the two words in the reCAPTCHA—"sample" or "Wills"—is from a scanned text that Google is attempting to transcribe for commercial purposes. After a sufficient number of people provide the same transcription for the same scanned image of the word, Google automatically updates its transcription of the source document with the website users' transcription.

24. Google profits in multiple ways from website users' transcription of scanned images through reCAPTCHA prompts. For example, third parties such as the New York Times pay Google to transcribe scanned images of archived versions of its newspapers, books, and other texts.

25. Google also profits from its own digitization of books, which it then offers in electronic format through Google Books.

26. Google also profits through advertising revenue it obtains on Google's various websites that make available to the general public Google's digitization of scanned texts, books, and Street View images. For example, the below reCAPTCHA example requires the user to enter two number sequences, the latter of which is "8001."



The image of "8001" appears to be a photo taken by a Google Streetview vehicle, such as the one pictured below, of an address on a house or other building.



After a sufficient number of users enter "8001," Google will know that the address of the building pictured is "8001," which Google's software then incorporates automatically into its Street View feature on Google Maps (a feature that allows Google Maps users to see images of what a person would see while driving on a particular street).

Google's reCAPTCHA Platform Is Deceptive and Unfair.

27. Although Google derives substantial profits from the transcriptions of images generated by users of websites that employ reCAPTCHA, Google does not disclose to website users that Google is profiting from their time and effort.

28. Google does not compensate website users for their time and effort in transcribing images.

29. Google, in essence, operates a highly profitable transcription business built upon free labor, which it deceptively and unfairly obtains from unwitting websites users, in violation of their rights as consumers, who are offered no choice but to transcribe words for the sole purpose of enriching Google.

Class Action Allegations

30. Plaintiff re-alleges and incorporates the allegations contained in the paragraphs above.

31. Plaintiff brings this action pursuant to Federal Rule of Civil Procedure 23 and Massachusetts General Laws chapter 93A section 9(2) on behalf of herself and a class, referred to herein as the "Massachusetts Class," consisting of:

All persons who have responded to a reCAPTCHA prompt in order to access a website or certain features of a website while they were Massachusetts residents.

32. Plaintiff brings this action pursuant to Federal Rule of Civil Procedure 23 on behalf of herself and a class, referred to herein as the "Nationwide Class," consisting of:

All persons who have responded to a reCAPTCHA prompt in order to access a website or certain features of a website while they were United States residents.

33. The Massachusetts Class and the Nationwide Class, together, are referred to herein as the "Classes."

34. All of the persons described in the class definitions in the three paragraphs above are referred to herein as “Class Members” or “Members of the Classes.” Plaintiff reserves the right to amend the definition of the Classes.

35. Because reCAPTCHA is the dominant platform in the CAPTCHA market, appearing on many of the most popular websites on the internet, including Google’s own Gmail service, many if not most internet users have responded to a reCAPTCHA prompt within the four years preceding this complaint. The Massachusetts Class totals at least hundreds of thousands of Massachusetts residents, and the Nationwide Class totals at least tens of millions of people. Accordingly, the Members of the Class are so numerous that joinder of all members is impractical.

36. Common questions of law and fact exist as to all Class Members and predominate over any questions solely affecting individual Class Members. Among the questions of law and fact common to the Classes are:

- (a) Whether, where, and when Google requires users to respond to a reCAPTCHA as a condition for accessing a website or a part of a website;
- (b) Whether Google discloses to website users that it profits from their time and effort in responding to reCAPTCHA prompts.
- (c) Whether Google compensates website users for their time and effort in transcribing images.
- (d) Whether Google was unjustly enriched by profiting from the free labor of website users without disclosure or compensation.
- (e) As to the Massachusetts class, whether Google violates Chapter 93A by profiting from the free labor of website users without disclosure or compensation.
- (e) The proper measure of damages.

37. Plaintiff's claims are typical of the claims of the Members of the Classes because, like Plaintiff, each Class Member responded to a reCAPTCHA prompt in order to access a website or certain features of a website.

38. Plaintiff will fairly and adequately protect the interests of the Members of the Classes and has retained counsel who have extensive experience prosecuting class actions and who, with Plaintiff, are fully capable of, and intent upon, vigorously pursuing this action. Plaintiff does not have any interest adverse to the Classes.

39. A class action is superior to all other available methods for the fair and efficient adjudication of this controversy. Furthermore, the damage that has been suffered by any particular Class Member is likely not substantial, and the expense and burden of individual litigation would make it impracticable for all Members of the Classes to individually redress the wrongs done to them. There will be no difficulty in the management of this action as a class action.

40. The Members of the Classes are identifiable through the records of Google and third-party website owners who employed the reCAPTCHA platform, and Plaintiff is a member of both Classes.

41. Google has acted on grounds generally applicable to the Classes with respect to the matters complained of herein, thereby making appropriate the relief sought herein with respect to the Classes as a whole.

COUNT I

(Violation of Massachusetts General Laws, Chapter 93A)

(Against Google on Behalf of the Massachusetts Class)

42. Plaintiff incorporates the foregoing paragraphs as if fully set forth herein.

43. At all relevant times, Google was engaged in trade or commerce within the Commonwealth of Massachusetts, including the trade or commerce of (i) offering its reCAPTCHA platform to websites with users located in Massachusetts; and (ii) using the reCAPTCHA platform on its own websites, including Gmail, which are widely used by website users in Massachusetts.

44. By operating a highly profitable transcription business built upon free labor, which it deceptively and unfairly obtains from unwitting websites users, Google has engaged in unfair and deceptive acts and practices in the conduct of trade or commerce in violation of Mass. Gen. Laws ch. 93A, § 2.

45. Google's violations of Mass. Gen. Laws ch. 93A as described herein were done willfully, knowingly, and in bad faith.

46. As a direct and proximate result of Google's conduct, Plaintiff and the Members of the Massachusetts Class have been injured by Google forcing Plaintiff and the Members of the Massachusetts Class to provide to Google, without compensation, their labor in deciphering the second image in a reCAPTCHA as a condition to accessing a website or a part of a website.

47. Plaintiff sent Google written demand for relief pursuant to Mass. Gen. Laws ch. 93A, § 9, identifying the claimants and reasonably describing the unfair acts or practices relied upon and the injuries suffered, on November 3, 2014. Google did not respond to the demand letter.

48. As a result of Google's willful, knowing and bad faith violation of Mass. Gen. Laws ch. 93A, Google is liable to Plaintiff and the Massachusetts Class for at least two and up to three times the damages that Plaintiff and the Massachusetts Class incurred, or at the very least the statutory minimum award of \$25 per each violation, i.e., each Massachusetts Class Member's response to a reCAPTCHA prompt, together with all related court costs, attorneys' fees, and interest.

COUNT II

(Unjust Enrichment)

(Against Google on Behalf of the Massachusetts Class and the Nationwide Class)

49. Plaintiff incorporates the foregoing paragraphs as if fully set forth herein.

50. Google received and retained wrongful benefits as a result of its forcing Plaintiff, the Massachusetts Class, and the Nationwide Class to provide to Google, without compensation, their labor in deciphering the second image in a reCAPTCHA as a condition to accessing a website or a part of a website.

51. Google's procurement and retention of the benefits of the labor of Plaintiff the Massachusetts Class, and the Nationwide Class under circumstances that make it inequitable for them to retain such benefits.

52. Google directly benefited from the conduct alleged herein through its retention of valuable transcriptions generated by members of the Massachusetts Class and the Nationwide Class, which Google obtained directly through website users' interactions with Google's reCAPTCHA platform.

53. As a result of the Google's wrongful conduct as alleged herein, it has been unjustly enriched at the expense of, and to the detriment of, Plaintiff, members of the Massachusetts Class, and members of the Nationwide Class. Under the common law doctrine of unjust enrichment, the benefits derived by Google through its reCAPTCHA platform rightfully belong to Plaintiff, the members of the Massachusetts Class, and the members of the Nationwide Class.

54. Plaintiff and members of the Classes have no adequate remedy at law.

Jury Demand

Plaintiff demands a trial by jury on all claims so triable.

Prayers for Relief

WHEREFORE, Plaintiff prays for relief in the form of an order as follows:

1. Allowing this action to proceed as a class/representative action under Mass. R. Civ. P. 23 and Mass. Gen. Laws ch. 93A, §9(2) ;
2. Awarding Plaintiff and Members of the Classes monetary damages, including double or treble damages, statutory damages, interest, and costs;
3. Awarding restitution (including through imposition of a constructive trust), including disgorgement of profits obtained by Google as a result of its of unjust enrichment;
4. Awarding counsel for the Plaintiff and the Classes their reasonable attorney's fees and expenses; and
5. Any other and further relief which the Court finds just and proper.

/s/ Thomas G. Shapiro
Thomas G. Shapiro (BBO# 454680)
Patrick J. Valley (BBO #663866)
SHAPIRO HABER & URMY LLP
Seaport East
Two Seaport Lane
Boston, MA 02210
Telephone: (617) 439-3939
tshapiro@shulaw.com
pvalley@shulaw.com

Counsel for Plaintiff

General Information

Court	United States District Court for the District of Massachusetts; United States District Court for the District of Massachusetts
Federal Nature of Suit	Personal Injury - Other[360]
Docket Number	3:15-cv-10160