2018 ENTRY FORM

(Note: word count 2,500)

Entry ID:	**02071**
Entry Title:	Re:scam - The World's Most Unhelpful Chatbot
Client:	Netsafe New Zealand
Product:	Re:scam
First Media Appearance Date:	6 November 2017
Category:	B - Social Marketing/Public Service
Category Description:	Marketing communications of a public service nature, including campaigns to promote social or behavioural change. This typically involves government department, local body or community service campaigns. Judges are looking for proof that your communications significantly contributed to a positive on-going social change, driving a valuable outcome of social good i.e. you changed how people think, what they do in line with stated campaign objectives.

Title:	Re:scam – The World's Most Unhelpful Chatbot
Client:	Netsafe New Zealand
Service:	Re:scam

1. Case Summary (0%)

Please write a brief summary of the case study and results not exceeding 90 words.

REQUEST FOR ASSISTANCE – STRICTLY CONFIDENTIAL

Dearest [Friend],

It is my pleasure to discretely write concerning a business matter of mutual benefit. My purpose of contacting you is to deliver a high volume of GOLD and SILVER, and because we are prohibited by the TOP OFFICIALS of the COMMERCIAL COMMUNICATIONS COUNCIL from doing so ourselves, consequently if you award us of this paper my colleagues and I will payment you the **total sum $11,000,000** upon delivery.

Please if you are willing indicate your interest in replying.

2. What was the challenge and what were the objectives? (10%)

What was the market context, what was the strategic challenge the client faced, what was the creative challenge the agency was set, and what were the short and long-term objectives that were set for the campaign?

"The real problem is not whether machines think - but whether we do"
-B.F. Skinner

Netsafe is a non-profit cyber-safety organisation that gives Kiwis the education and tools to keep them safe online; from online bullying to healthy social media usage.

But there was one type of digital danger that hadn't yet been tackled at scale.

Phishing attacks (fraudulent attempts to obtain personal information)[1] have increased by 65% worldwide since 2015.[2] Closer to home, in the past year New Zealanders lost at least $257m to cyber crime.[3] That's more than the annual revenue of TradeMe[4], or a shade under 170 houses in Grey Lynn.[5]

Sadly, the real losses are much higher. The above figure only takes into account **reported** losses. It's not hard to see why most victims stay silent. It's embarrassing to admit to being fooled. While it might be tempting to look with scorn at those who have fallen for scams, 30% of us open these emails, and 12% of us even click on the malicious attachment or link.[6]

This global criminal business was only growing more dangerous and sophisticated. Phishing rates have increased across almost all industries[7], with nearly 1.5m new sites created every month.[8] They are nearly impossible to prosecute, and victims have almost no avenues of recourse.

Email scamming was a widespread issue that was genuinely ruining lives, but we still felt invincible online.

The first and only line of defence is not to get scammed in the first place, by being aware of the red flags and avoiding them.[9]

---

[1] In *Stamp, Mark & Stavroulakis, Peter*. Handbook of Information and Communication Security, Springer, ISBN 978-3-642-04117-4.
[2] Anti-Phishing Working Group – 2018: http://www.antiphishing.org/apwg-news-center/
[3] Amy Adams, Ministerial Address to 2016 Cyber Security Summit: https://www.beehive.govt.nz/speech/ministerial-address-2016-cyber-security-summit
[4] Deloitte Top 200: https://www.top200.co.nz/wp-content/uploads/Top200-Awards-2017-Rankings.pdf
[5] Barfoot Market Report: https://www.barfoot.co.nz/market-reports/2018/january/suburb-report
[6] Verizon Data Breach Investigations Report (DBIR) 2018: www.verizonenterprise.com/verizon-insights-lab/dbir/2017/
[7] Symantec – June 2017 Spam & Phishing Intelligence Report: https://www.symantec.com/connect/blogs/latest-intelligence-june-2017
[8] Webroot Cyber Security: https://www.webroot.com/us/en/about/press-room/releases/nearly-15-million-new-phishing-sites
[9] Microsoft – Cyber Crime in New Zealand – What You Need To Know: https://blogs.business.microsoft.com/en-nz/2017/11/09/infographic-cyber-crime-new-zealand-need-know/

Netsafe's **objectives** for the campaign were threefold:

1. **Make people aware of the dangers:** educate internet users on phishing scams. Measured by earned media coverage (1.5m Kiwis and 10m globally)
2. **Give internet users a tool to fight back against phishing scams:** measured by engagement with our campaign (more than 20,000 actions taken)
3. **Make people aware of Netsafe's role in keeping Kiwis safe from harm online:** by doing so, raise the profile of Netsafe: measured by attribution (95% with 2+ key messages) and an increase in website traffic (30,000 unique browsers)

3. What was the strategic thinking that inspired your big idea? (15%)

What was the insight or insights identified as key to unlocking the solution? How and why did the strategic thinking address the objectives set?

THINKING OUTSIDE THE (IN)BOX

"All the world is made of faith, and trust, and pixie dust."
-J.M. Barrie, Peter Pan

1. To trust is human

Human beings are naturally predisposed to trust. It's in our genes as a survival mechanism that has served our species well.[10] People simply don't have the time or capacity to calculate the risks that come with every choice. We are guided by our gut and consistently make hot-headed, emotional decisions. Even with – especially with – our money.[11]

Kiwis in particular are trusting animals.[12] NZ is known for its low levels of corruption, but as soon as we enter digital domains we become connected to some of the most untrustworthy parts of the globe.

2. Scammers play off our trust to lure in their victims

Unfortunately, this means we can be manipulated. From "Microsoft tech support" to Bernie Madoff, fraudsters play off humanity's trusting nature to fool their marks into a false sense of security by appearing to be the real thing. It's not just the elderly – millennials are the most likely victims.[13] Nobody ever thinks it'll happen to them. And nobody ever sees it coming.

3. Strategic Approach: Turn the tables on scammers by using their very own psychological tools against them

Phishing scams have been around as long as email. They were not new news.

With a massive budget it might have been possible to reach a complacent audience, but without one we had to make phishing worth talking about in a non-tokenistic way, by giving the public the power to make a difference, and a reason to share their experience to drive awareness for us.

[10] Harvard Business Review – Rethinking Trust: https://hbr.org/2009/06/rethinking-trust
[11] Quartz – Humans are born irrational: https://qz.com/922924/humans-werent-designed-to-be-rational-and-we-are-better-thinkers-for-it/
[12] NZ Herald / PWC – Kiwis are too trusting: https://www.nzherald.co.nz/co-curated-content/news/article.cfm?c_id=1504362&objectid=11943583
[13] Norton Cyber Security Insights Report 2017: https://www.netsafe.org.nz/2017-norton-cyber-security-insights-report/

Scam-baiting had been done brilliantly in the past by comedians such as James Veitch, but always at a one-on-one level – wasting your own time in the process.

We needed a similarly **personal connection**, but we also needed to do this at **scale**.

Connecting these two dots led us to our solution: **chatbots**

Predicted to be involved in 85% of customer service interactions by the year 2020[14], their rise looks inexorable. They are already blending in. Last year 27% of consumers weren't sure whether their last customer service interaction was with a human or an AI.[15] They're as deceptive as scammers - and just as far-reaching.

What if we turned this ambiguity on the scammers themselves?

Strategic proposition:

Use the scale and personality of chatbots to beat scammers at their own game.

4. What was your big idea? (10%)

State in one sentence. What was the core idea that drove your effort? Consider 'idea' in the broadest sense, i.e. ranging from communication-based to the creation of a new service or resource. The idea should not be your execution or tagline.

An AI-powered chatbot that imitated human victims, wasting scammers' time and protecting real people from harm.

[14] Forbes / Gartner - Chatbots and the Customer Experience: https://www.forbes.com/sites/blakemorgan/2017/03/21/how-chatbots-will-transform-customer-experience-an-infographic/#5ca4b61d7fb4

[15] PWC – Bot.Me: A revolutionary partnership: https://www.pwc.com/us/en/industry/entertainment-media/publications/consumer-intelligence-series/assets/pwc-botme-booklet.pdf

5. What was the creative execution and how did it bring the big idea to life? (15%)

Describe the creative work that delivered the big idea.

"Stop communicating products and start making communication products."
-Gareth Kay

THE KING-PHISHER

Re:scam was an AI-based initiative that gave people a tool to fight back against scammers.

When someone received a phishing email, they could forward it to me@rescam.org. Our program then picked up the conversation and replied to the scammer based on the email. Replies were designed to lead scammers on for as long as possible with exchanges that wasted limitless hours of their time.

If scammers were busy talking to a robot, they weren't talking to real people.

This was a good first step, but at its heart Re:scam was a faceless entity, not built to be shared *en masse*. Because we had no media budget, if we wanted to give ourselves a chance of breaking into culture and driving mass awareness we needed to give it some personality.

Or rather, multiple personalities.

MEET THE BOTS

We introduced AI cat-phishing to the world with a deliberate blend of human and computer-generated creativity.

We engaged IBM's AI '*Watson*' to help analyse the content of messages and formulate responses, and created a digital video as the centre-piece of our communications. This mirrored the multiple personalities of Re:scam by showing different C.G. faces and voices flickering in and out.

To show that anyone could be a victim of an email scam, Re:scam was created to mimic various types of personalities. With deliberate spelling mistakes and malapropisms, each "character" had their own backstory and unique way of talking.

From the retiree asking "*The Illuminati*" if they had a bingo night he could join (and who sent his bank details through One. Number. At. A. Time), to the single mother who was excited to win big money, each was programmed to be as frustrating and time-consuming as possible, while remaining human enough to avoid detection. Sometimes our bots would accuse the scammers themselves of being bots.

Every time they got a response, they now had to second guess themselves.

6. What was the communications strategy? (10%)

Outline the media and communications thinking and strategy that brought the creative solution to life in the most powerful and relevant way for the target audience.

The success of this campaign hinged entirely on securing earned media coverage to drive users to the website and spark conversations about phishing.

We knew that - like phishing scams themselves - this idea was so big it could become borderless. We therefore designed our communication strategy to make Re:scam as global as possible. This would have the added bonus of making us "*world famous in New Zealand.*"

To overcome our lack of media budget, the PR strategy was modelled on other successful tech launches. We offered an exclusive to *The Project* and seeded the campaign on Reddit. This secured a user base to provide immediate engagement with the technology.

We encouraged journalists to try out the bot for themselves,[16] giving them behind-the-scenes access to bolster its credibility.

The site itself then became the ultimate interactive media kit to make storytelling as easy as possible.

The Re:scam avatar talked users through the product; hilarious conversations were hosted on-site which were then embedded on media stories; and an FAQ section ensured that time-poor journalists had all they needed to file a story quickly.

Using this we pitched and secured coverage from the global media.

Online recruitment videos, social media, digital billboards and Re:scam's PR coverage (all earned) encouraged people to forward their scam emails.

Re:scam was built to chat, so it was a natural in social. A Twitter account (appearing to be run by the bots) posted the best exchanges and alerted our audience to trending scams.

[16] Stuff.co.nz – Why I Posed As A Russian Bride (It's Not What You Think): https://www.stuff.co.nz/technology/98779967/why-i-posed-as-a-russian-bride-to-trick-a-scambot

List all consumer communications touch points used in this campaign.

- Online video
- Digital (website)
- Digital billboards
- PR
- Social media (Twitter, Facebook, Reddit, Imgur)

7. What was the $ spend? (0%)
Outline the media and production spend on the campaign. Use actual spend rather than rate card. In the case of donated media please list the rate card value separately from the bought media spend.

Media Spend:	N/A

Outline the media spend in relation to competition and versus last year:

As the only NZ non-profit that focuses on online safety, Netsafe has no direct competition. (Apart from the scammers themselves).

Creative Production Spend:	N/A

8. What other marketing efforts were used in conjunction with this campaign? (0%)
List all other marketing or communications programmes not considered part of this campaign, that also affected the results e.g. coupons, sales promotion, planned PR, sampling, direct response, point-of-purchase, etc.
Indicate the extent to which any revised pricing, distribution or promotion programmes also affected the results.
Any marketing communications that contributed significantly to delivering an integrated campaign strategy and results should be described elsewhere in the entry form and any relevant contributing partners acknowledged in credits separate to the entry form.

There were no other marketing communications from Netsafe NZ at the time of this campaign.

9. What were the results? (40%)
Outline the results achieved by the campaign against the short and long-term objectives set, provide conclusive proof that it was the campaign that drove the results and outline the return on investment.
In this section, the judges will be looking to see a clear cause and effect between the communication activity and business performance over-time. Show the compelling evidence that will convince even the most cynical finance director. They will be awarding points on the following basis:

Overall achievement against objectives (10%)
Clear demonstration of long term results beyond 6 months (5%)

Convincing proof that the results were a direct consequence of your campaign, the inarguable evidence. (15%)
Return on investment. This could be expressed as an ROI (Return on Investment) figure, or some other numerical way of demonstrating commercial payback that justifies the investment in the campaign in the first place (10%)

This campaign delivered results that would make any marketer green with envy. We won the lottery in two South American countries and began long-distance relationships with several Russian supermodels. We were inducted into the Indonesian Police Force and became reacquainted with countless long-lost and fabulously wealthy distant relatives. To top it all off, we are soon to take possession of several crates of gold bullion - as soon as our payment for shipping clears.

Real results were almost this good.

1. **<u>Make people aware of the dangers:</u>** educate internet users on phishing scams. Measured by earned media coverage (1.5m Kiwis and 10m globally)

Our campaign immediately took on a life of its own. In the hours following launch, Re:scam received more than 3,600 emails and more than 18,000 unique browsers.

The NZ media fell over themselves to cover it, with a reach of **4 million+** across all major networks.

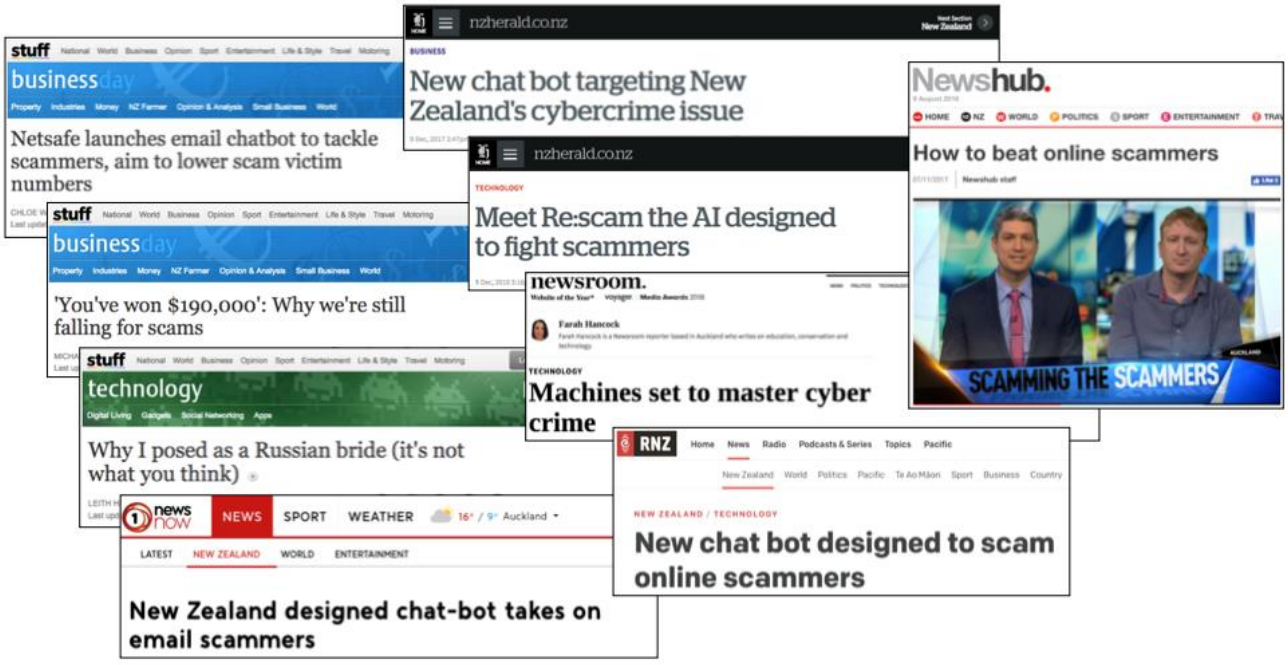


Fig 1: Online local news articles about the campaign on Stuff.co.nz, New Zealand Herald, RNZ, One News, Newsroom, Newshub

Followed by eye-watering global reach:

This was in excess of **300 million+** over the course of the campaign, as Re:scam trended on the likes of The BBC, The Guardian, The Daily Mail, El Pais, and many more:[17]

Fig 2: Online global news articles about the campaign from The Guardian, Independent, El Pais, VICE, BBC, et al.

Our digital video campaign centre-piece achieved **over 4.5 million** views across Facebook and YouTube.

[17] Total of average daily audience per media outlet that ran a story on Re:scam

During the campaign period, **273,000 unique browsers** visited the site, with the vast majority of this driven by earned media. In total, the website achieved over **400,000 page views,** with users spending almost a minute and a half on the site on average.[18]

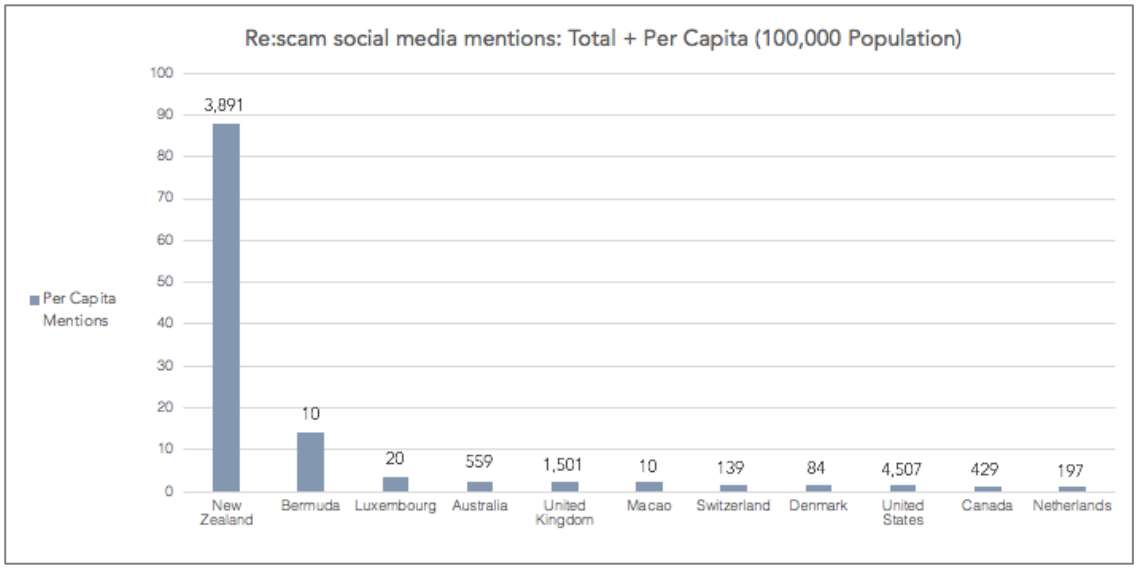
Once people were aware of Re:scam, they couldn't help but dive in to start scam-baiting and sharing their results. After all, schadenfreude is one of life's guilty pleasures.

2. **Give internet users a tool to fight back against phishing scams:** measured by engagement with our campaign (more than 20,000 actions taken).

Our audience relished the sheer joy of revenge served cold:

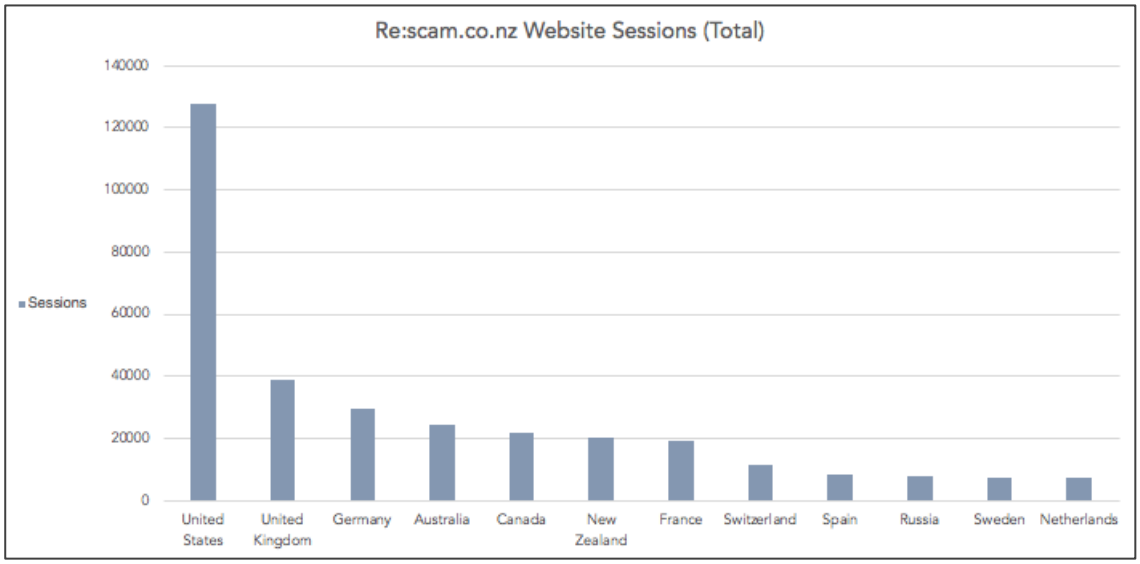
Over the campaign period of two months, a staggering **210,000 scam emails** were forwarded to us from every corner of the globe.

New Zealanders led the charge, being far more likely than others to mention Re:scam in social media:[19]



Kiwis were also **much more likely to engage** with the Re:scam site, coming in 6th overall and 1st per capita:[20]

[18] Re:scam website data
[19] Netbase social listening
[20] Google Analytics data: Re:scam website



As soon as people started forwarding their emails, our chatbot then got to work. Over **1 million emails** were sent to scammers.

In total, this **wasted 5 years of their collective time**.[21] Every minute of which was time that they couldn't spend attacking real victims.

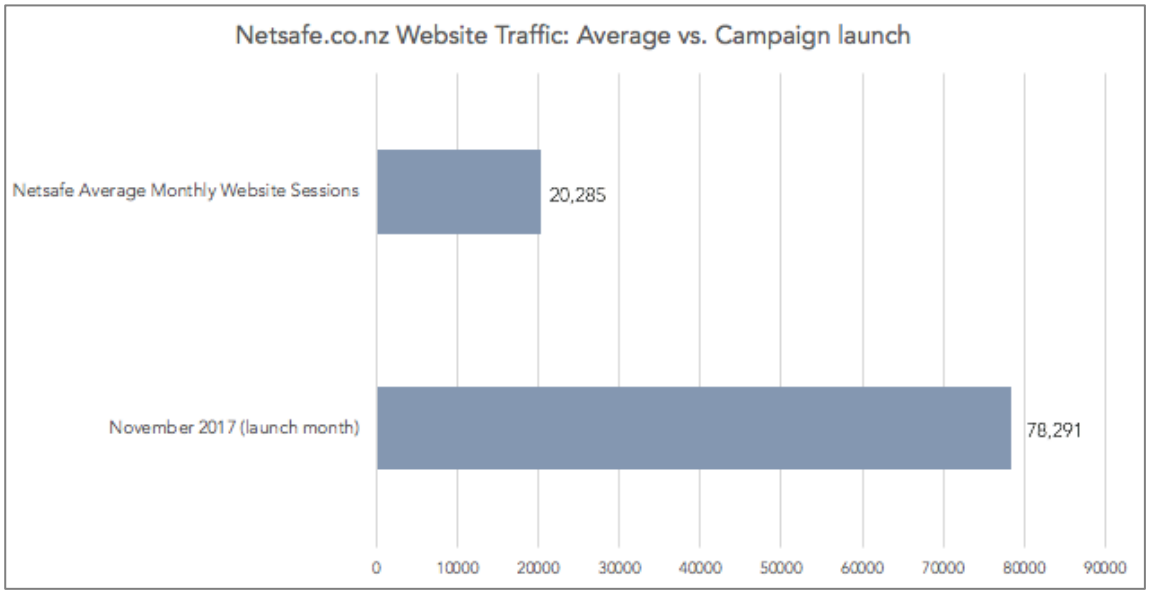
At one point during the campaign there were 1,000 concurrent conversations happening with scammers around the world. The longest involved over 170 exchanges.

All of this data is available for cyber-crime agencies including the NZ Police and Interpol.

2. **Make people aware of Netsafe's role in keeping Kiwis safe from harm online:** by doing so, raise the profile of Netsafe: measured by attribution (95% with 2+ key messages) and an increase in website traffic (30,000 unique browsers)

We also drove awareness of Netsafe itself, not just the Re:scam campaign. Site traffic to Netsafe's own website quadrupled over the Re:scam launch month to nearly 80,000 unique users:

[21] This is a conservative estimate based on average human typing speed and number of responses.



We successfully raised awareness of Netsafe's work in cyber-security. Every major piece of media coverage named Netsafe with 2+ key messages.

As a result of this campaign Netsafe has received validation from industry bodies - winning the top prize at the Southern Hemisphere's only anti-fraud awards for "*capturing the attention of the world*."[22]

Netsafe has also been given a profile boost by being invited to more speaking engagements, including being invited to speak at *AI Day* alongside Rocket Lab.

All of this was achieved with no other Netsafe communications in market: these results could only have been a direct result of the campaign, despite our limited budget.

<u>Evaluating ROI</u>

This campaign was never about securing a financial return on investment for Netsafe (a non-profit). It was about using a big idea to drive awareness and education of email phishing, which we clearly achieved as per the above results.

Even so, let's quantify things. Social listening showed over 230m potential impressions through online conversation.[23] In addition, the campaign had a reach of more than 300m people through media articles, with more than 400 stories in over 20 languages.[24]

[22] Ian Tuke, Chair of NZ Fraud Film Festival, http://www.scoop.co.nz/stories/CU1803/S00112/netsafes-rescam-chat-bot-takes-home-anti-fraud-award.htm

[23] Netbase social listening analytics, data gathered during 2 month campaign period of Nov – Dec 2017. This does not include any Facebook data – a single "LADBible" post on Re:scam achieved 2.4m video views, 21k Reactions and 7.7k Shares – not counted here.

[24] Total of average daily audience per media outlet than ran the story

The earned media value of this coverage is estimated to exceed $30m, delivering a PR Media Value of 1:1,617 ROI.[25]

Secondly, while we can't make absolute claims on how much money Re:scam saved potential victims or publicly disclose expenditure, we can look at an estimate of the average scam loss, combined with the number of individual conversations we were having with scammers instead of real victims. [N.B. Internal data removed for publication].

If we assume that just 0.01% of our conversations with scammers might have saved a real victim from harm, based on our calculations we are left with an ROI of 1:30.

Finally, Netsafe was contacted by one New Zealander who, thanks to our campaign, was now aware that he was about to become the victim of a Russian romance scam. Saving even one person from the hardship of being scammed has made the whole thing worth it.

Resulting in thousands of unhappy scammers:

"I HAVE RECEIVE MORE THAN 85 EMAILS FROM YOU AND I HAVE ALSO SEND 85 EMAILS TO YOU AS WELL WITHOUT NO WAY FORWARD I BELIEVE YOU ARE NOT SERIOUS AT ALL..."[26]

And one happy client:

"Because of Re:scam, Netsafe was offered opportunities to participate in new conversations, [from] AI conferences [to] brand experience conferences. After 20 years in the industry, the invitations to speak at these events for the first time is a clear indication that [Re:scam] has positioned Netsafe as a thought-leader in industries that relate so closely to online safety."

-Sean Lyons, Director of Education and Engagement, Netsafe

TOTAL WORD COUNT (count only words you insert in answer boxes 1 - 9): 2433

[25] Calculated as: advertising equivalent value of each story x 3 = PR media value

[26] Transcript of scammer email exchange via Re:scam bot