mimecast

The Size and Shape of Workforce Risk

A CYENTIA INSTITUTE STUDY BASED ON DATA FROM ELEVATE SECURITY

Introduction

Everyone at your organization has a different role. From the maintenance staff to the CEO, the skills required to keep the gears of an organization's bureaucracy running smoothly differ widely.

Even though people vary in background and duties, security controls for workforce risk are often universally and indiscriminately applied. Policies govern appropriate use for everyone, all are subjected to the same security awareness training course, phishing simulation emails are sent to every address, and everyone's traffic runs through the same set of network appliances and anomaly detectors. This one-size-fits all approach has its advantages. It is certainly easier to deploy in an organization, and provides a sort of universal, minimal set of mitigations.

But if a universal approach worked, we wouldn't be here, would we? Some users are security pros who are fastidiously cautious in all their online interactions. For these users, the many security guardrails organizations put up won't reduce their already low risk or lower the risk for the organization as a whole. On the other end of the spectrum are the users that the CISO has nightmares about. Users who click every link that pops up on their screen and download every attachment while streaming from illegal media sites. All the policy and training likely won't put a dent in the outsized risk those users represent. Moreover, despite the desire to have a universal approach, many organizations vary in their ability to deploy changes uniformly.

"

On the other end of the spectrum are the users that the CISO has nightmares about. Users who click every link that pops up on their screen and download every attachment while streaming from illegal media sites.

This raises a bunch of pretty pressing questions:

- **How do we identify the riskiest users?**
- **What exactly do we mean by "risky" anyway?**
- **Are there a lot of risky users or just a few?**
- **Are some users more prone to be risky than others?**
- **How can we help them engage in less risky behavior?**
- **What kind of impact on our incidents could proactively identifying our risky users have?**

In this report we start to answer some of these questions. In particular we leverage data from Elevate Security to clarify what exactly makes a 'risky user'. We'll present some concrete numbers so you can understand how your users or departments measure up to everyone else. Finally, we'll see that some types of risky behavior are likely indicative of other kinds.

Materials and Methods

The analysis in this report was, unless otherwise cited, conducted on data provided to the Cyentia Institute by Elevate Security. The data include 15.1m unique events associated with 168k users spread across more than 3.8k organizational departments and aggregated in Elevate Security's data platform. The data include events starting in January 2016 through December 2021. Not all users or departments measure all event types. Any analysis below is conducted only on those departments/users who have measurable different event types.

Key Findings

Just a small handful of users create most security events.ww

Many users are low risk!

- 76% have never clicked a phishing email.
- 93% have never had a malware incident.
- More than half won't have secure browsing incidents.

But some users are nightmares

- 4% of users are responsible for 80% of phishing incidents, some clicking as often as twice a month.
- 3% of users are responsible for 92% of malware events, 1% will average an incident every other week.
- 12% of users are responsible for 71% of secure browsing incidents, 1% will trigger 200 events per week.

Controls can help mitigate risk but are variable

- 17% of departments block no malware.
- Most departments block the majority (>95%) of phishing or less than <10%, there is little in between.

Risk leads to risk. High risk users tend to be high risk in multiple ways.

Pinpointing Risky Users

So why focus on users at all? After all it's the incidents we care about. We like to say around here "Cyber security is human security"[1], because a large fraction of security incidents have a human in the loop moving that mouse and interacting with that questionable software.

In fact those who read our prior report may recall that over two-thirds of data breaches and nigh 90% of losses from the most damaging cyber incidents of the last several years tie back to employees doing things (mostly unintentionally) we don't want them to do. Risky actions contributing to those numbers include phishing schemes (~37% of all breaches), malware (80% of system intrusions), and infections from browsing to malicious sites (37% of malware breaches).[2]

So now that we've established the groundbreaking revelation that "people cause security incidents", a reasonable follow-up is "which people?". In the remainder of this report, we are going to define exactly what a "risky" user is in three different categories: phishing, malware, and secure browsing.

Phishing Events

TL;DR:

- Some users get many more phishing emails than others (100s per year vs. a few).
- The more emails a department gets the better they are at blocking them.
- Most users won't click the emails that do make it to their inboxes.
- But some of those who do will click a lot (as much as one click per week).
- Subjecting all users to the same level/type of treatment is counterproductive.

According to the 2021 Data Breach Investigations Report (DBIR), phishing is the #1 threat action used in data breaches and #2 among all reported security incidents. Clearly, reducing the propensity of users to click on phishing emails would go a long way toward reducing cyber risk. Let's examine what the flow of phishing emails to users looks like, how organizations attempt to stem the tide of those tempting links, and how often users are likely to slip up. In our last report, we focused on simulated phishing exercises, this year we are blessed with a bounty of real phishing emails. What follows is analysis on real live attacker-initiated phishing attacks.

1 Huh that's a pretty good title. I wonder if there is a talk about... https://www.youtube.com/watch?v=Xfo6IWH6B-k

2 Statistics sourced from the 2021 and 2020 Verizon Data Breach Investigation Reports

Frequency of phishing attempts

Before we even get to asking about what percentage of phishing emails get a click from users, it's worth asking exactly how many actually get sent to users on a yearly basis. Let's see if we can narrow this down to something slightly more specific than "a lot". Figure 1 is a "quantile dot plot" showing exactly the rate of phishing emails (as measured in emails received per user per year) sent to both departments and users.

Each dot represents 1% of users (or departments) in our sample. Wider areas where the dots are stacked indicate a delivery rate is more common while when the dots are off on their lonesome it means that not many users or organization's inboxes are being filled at that rate.

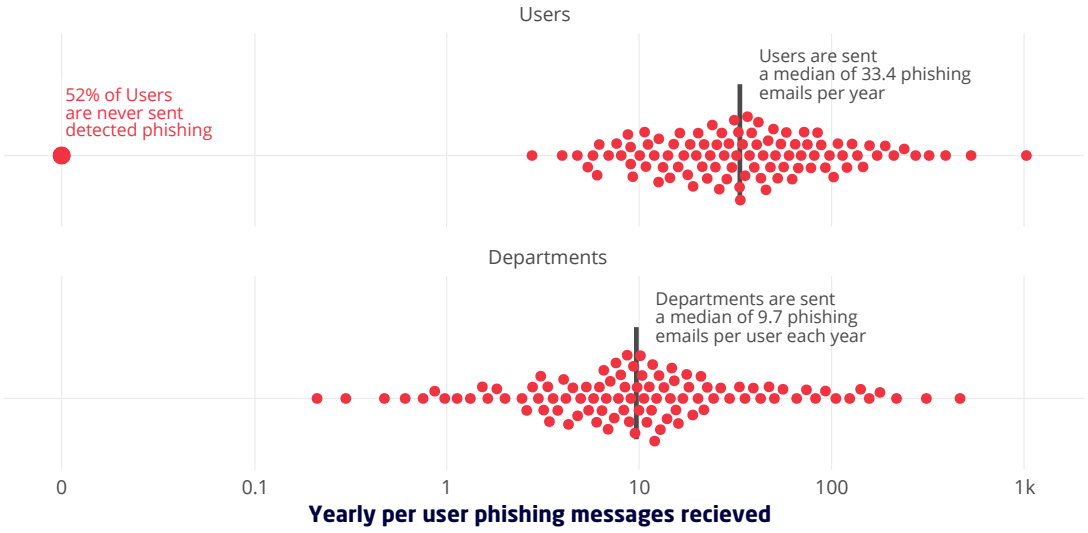

Figure 1: Number of phishing emails received per user per year

What do all those dots mean for our users? A little over half of users never receive phishing emails, but among those that do, around 30 per year is typical, a little less than three a month. Eagle-eyed readers will note that Figure 1 is 'log scaled' meaning each tick represents an order of magnitude increase. This means that 4.7% of users are sent (on average) one for every working day of the year. But it also means that many users receive as few as three a year, and in fact, 51% have never been sent a detected phishing email in their tenure.

The fact that a little over half of users don't see phishing emails emphasizes the importance of "pinpointing" more at-risk users for various interventions. While those users should certainly be aware of the dangers of phishing, those at the helm of the email addresses receiving phishing emails every day are the ones who need to be trained to be vigilant and/or equipped better.

There is similarly wide distribution among departments.[3] Interestingly, no departments in our sample were immune from phishing emails in the same way about half of users were. Those "phishing zero" users are well distributed among departments, pulling down the overall distribution, with a typical department receiving about 10 emails per user per year. Knowing where your department lies along this spectrum can help you justify and prioritize both automated and human interventions when it comes to phishing.

3 It's worth asking, "why do some users get so many phishing emails?", and one we want to explore in the future. For now, we want to use this report to give people a baseline for how much phishing is out there.

Block rate for phishing events

Speaking of those automated anti-phishing controls, of course we know that not every phishing email sent to users actually makes it to their inbox. Companies invest heavily in anti-spam and anti-phishing technology to keep themselves safe. So how effective are these controls? We take a look in Figure 2.

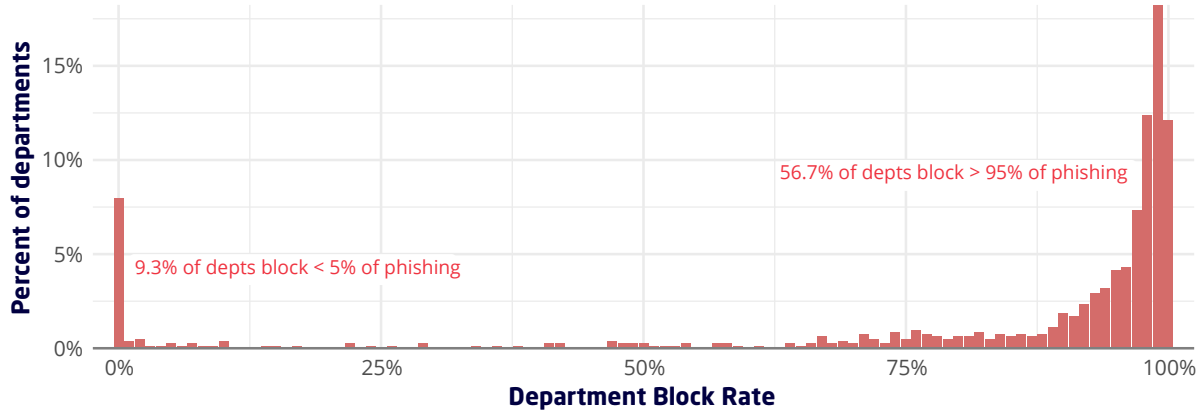

Figure 2: Proportion of phishing messages blocked by department

Department level block rates run the gamut of efficacy[4], but concentrate between 75% and 100% with more than half blocking 95% or more. That suggests most organizations are pretty aggressive in their attempts to catch phish before they spawn trouble, and it appears those efforts are effective. It's also instructive to look at whether departments which are receiving a lot of phishing are good at blocking it and, glancing at Figure 3, practice does appear to make things closer to perfect.

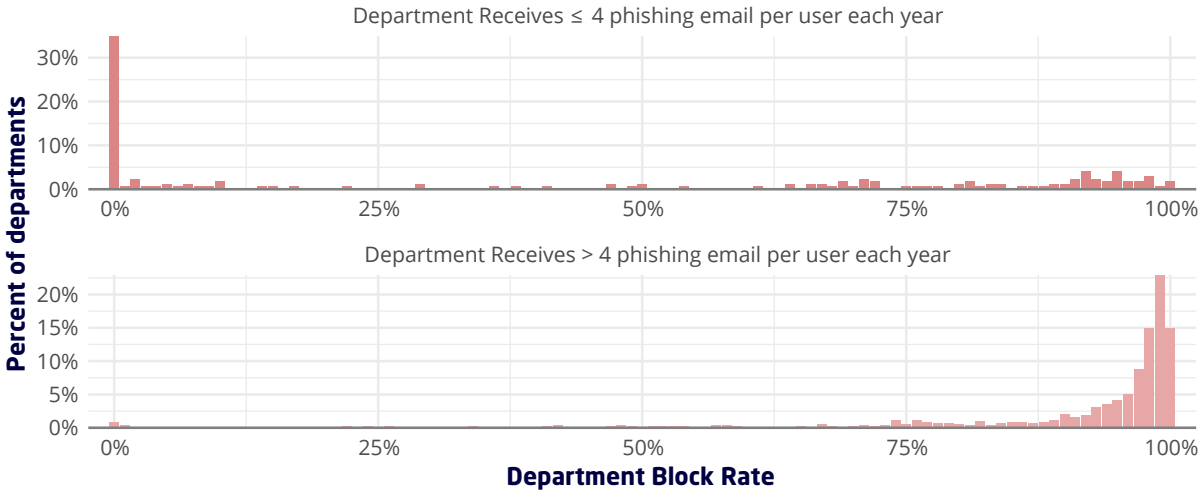

Figure 3: Department block rate for departments receiving different volumes of phishing emails

4 In general, we find that block rates vary quite a bit by department even within a single organization. We'll see this phenomenon repeated with malware, but for now we'll just say that blocking doesn't seem to be a uniformly effective control within organizations.

Figure 3 shows that departments that receive less than 4 phishing emails per user per year (the bottom quartile, shown in blue in Figure 3), are much more likely to not block any email that come in, meanwhile the top 3 quartiles are much more likely to have an effective blocking program (red histogram). This is almost certainly borne of necessity; without effective controls, email would likely stop being useful to most users.

Click rate for phishing messages

No matter how well your email filters are humming along, some of those messages are bound to get through. And given that organizations might have 10s or 100s of thousands of employees a substantial volume are likely to get through. So, when a user gets a notification about a phish-y email what is the likely outcome? Figure 4 has some good news for us about user click rates of unblocked phishing emails.

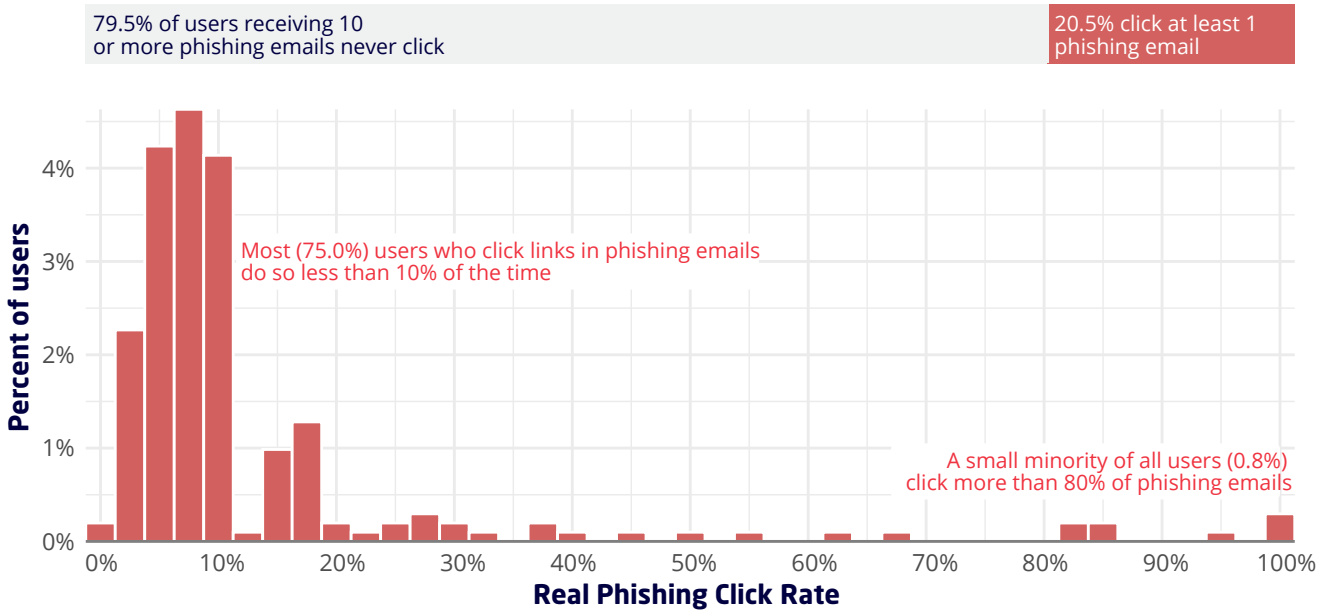

Figure 4: User click rate of phishing emails

The bar at the top reveals a roughly 80-20 split among users. Thankfully, the scales are tipped toward those that never take the bait (at least, not during the period of time we measured). Even among those users who do "take the bait" most (75%) do so less than 10% of the time[5]. Unfortunately, a small proportion of users swallow most of the lures tossed their way hook, line, and sinker. So, what percent of trouble is that small percentage of users causing? Well, it turns out quite a bit.

5 A quick data note: We are filtering down to users who have received at least 10 phishing emails so we have at least a little bit of statistical certainty around their click rate.

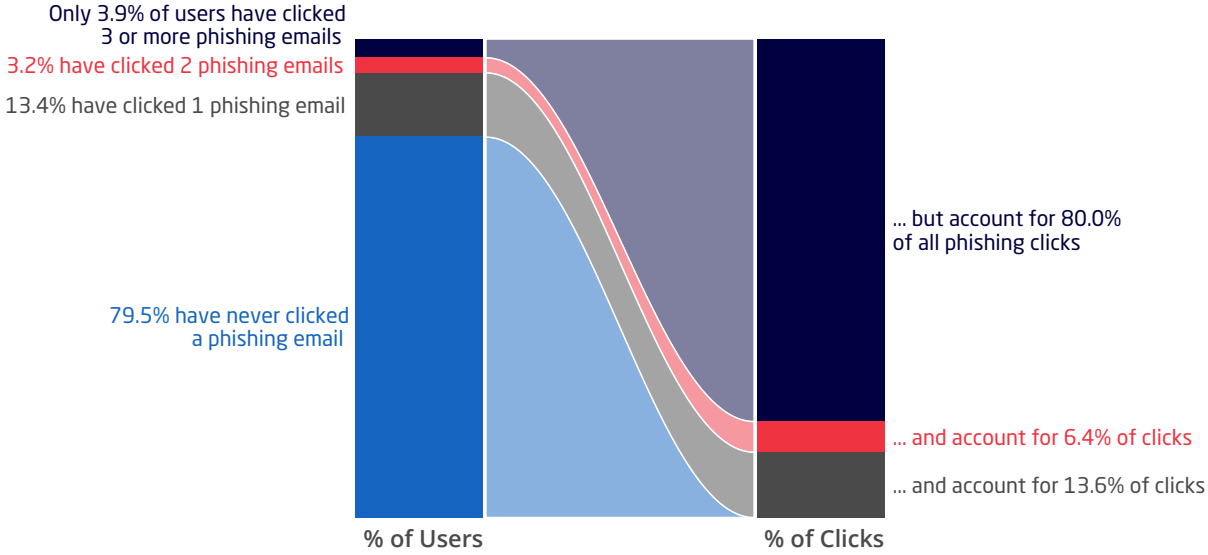

Figure 5: Phishing clicks by users

Figure 5 shows that it's a little bit worse than just the 80/20 rule for phishing clicks, it's more like the 80/4 rule. That is 4% of users are responsible for 80% of the actual clicks of phishing emails that occur.

These three measures on phishing alone show the variability in risk and the value of being able to measure it. A department receiving a phishing email per day per user might seem like a huge risk, but if good technologies are in place to block the vast majority of those emails and the users in that department are particularly adept at avoiding the ones that slip through, that department may be less risky than one that receives fewer but blocks none and is a little more click happy.

"

Figure 5 shows that it's a little bit worse than just the 80/20 rule for phishing clicks, it's more like the 80/4 rule. That is 4% of users are responsible for 80% of the actual clicks of phishing emails that occur.

Expected frequency of successful phish

So, let's try to combine everything we've learned above to get an idea of how many phishing events a user is likely to experience each year. To do this we built a probabilistic model based on the above observations about the rate of phishing emails sent, the apparent block rate, and the click rate.[6] The result is Figure 6.

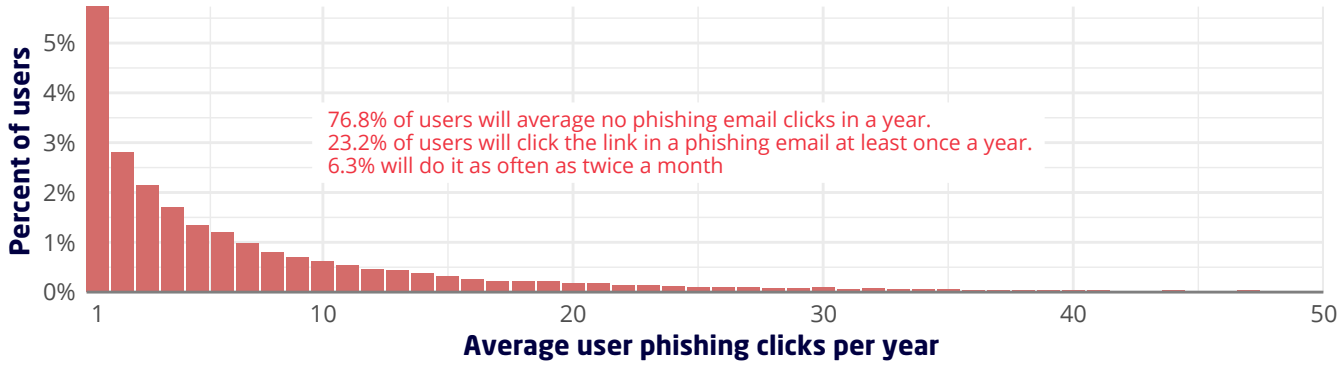

Figure 6: User phishing clicks per year

This figure will hopefully give organizations actionable ideas about exactly who is a high-risk user and who isn't. In particular,

- **22.9% of users will actually click through one phishing email per year, among those:**
  - **75% will click about once 1 a year**
  - **50% will click more than 3 a year**
  - **25% will click almost 9 a year**
  - **1% will click more than 52 a year (~ ONCE A WEEK)**

Those 1% of users are the result of those unfortunate souls that receive many, many phishing emails, whose departments aren't great at blocking them, and their average click rate is high. Those are the users we need to keep an eye on.

6 Specifically, we fit a lognormal distribution to the rate of emails received by user, combined that with a beta regression model predicting the block rate from the correlation between rate of emails received and block rate, and finally assigning users a click rate from a beta distribution fit to the data in figure 4.

Malware Events

TL;DR:

- Similar to phishing, some users are going to encounter much more malware than others.
- The vast majority will not experience any.
- Anti-malware controls don't seem to be deployed uniformly even with an organization.
- Most departments won't experience any malware events on a yearly basis, some will experience it weekly.

Compromised credentials are bad. But you know what's also bad? Malware. Arguably as old as personal computing itself, attackers have continually used malware to wreak havoc across organizational networks. The 2021 DBIR found that malware was the most common threat vector used in the middle and latter phases of the chain of events leading to data breaches. Of course, this makes sense, as malware is a virtual Swiss Army knife for elevating privileges, maintaining access, broadening control, capturing data, and accomplishing all manner of nefarious goals.

Similar to how we examined phishing across users and departments we wanted to do the same with malware. Dig in and see what constitutes a "high risk" and look into some of the ways organizations are attempting to address malware. One challenge here is that organizations' approaches to malware are nearly as diverse as malware itself. Our malware events are broken into three rough categories: "Blocked", "Downloaded", and "Executed". "Blocked" indicates it never made it onto the intended machine, "Downloaded" means it was downloaded but never had a chance to execute, and of course "Executed" indicates that some poor user was tricked into running the badness.

"

Similar to how we examined phishing across users and departments we wanted to do the same with malware ... One challenge here is that organizations' approaches to malware are nearly as diverse as malware itself.

Frequency of malware encounters

Just like Figure 1 focused on how many phishing emails got sent to a user, Figure 7 measures how many malware events occurred for each user and department. It provides a somewhat more hopeful picture. In general users are very unlikely to experience any type of malware event, with 94% of users and 57% of departments never recording a malware event.

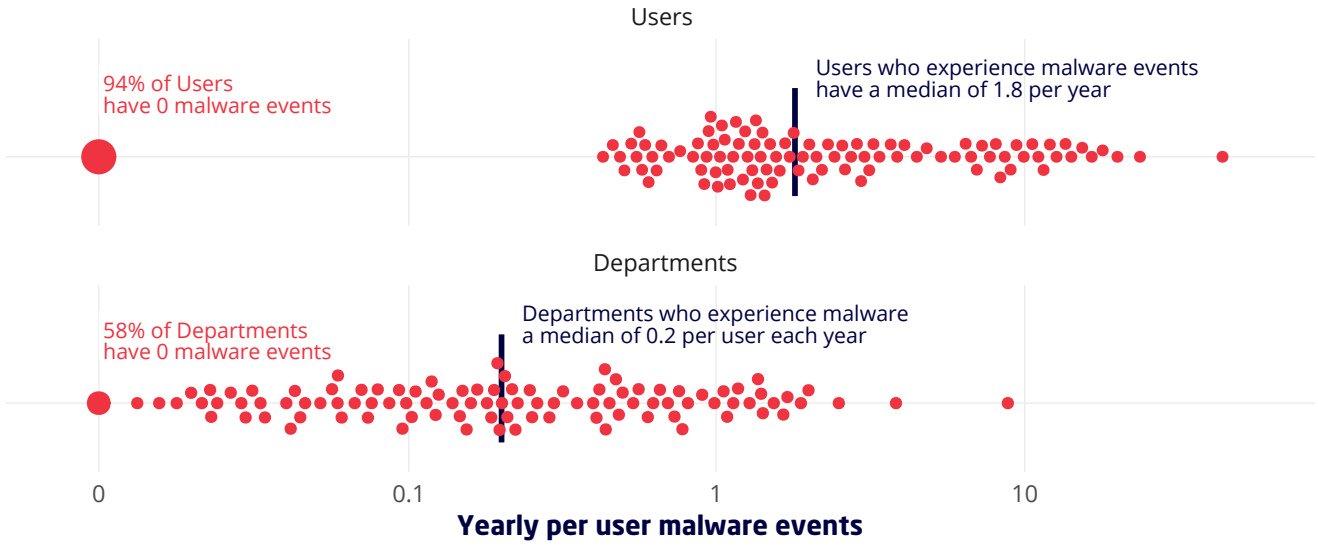

Figure 7: User and department malware events per year

Even those users who do experience malware do so at extremely low rates, averaging around one event per year. Once those "0 event" users get scattered among their respective departments and we calculate average rates for departments, we see that departments typically have about one event every three years per user. Somewhat comforting, and that may be due to more anti-malware controls separating malicious code from end users. What is not comforting is the same heavy tail of this distribution. 10% users average more than 11 events per year, with 1% as high as 27. These are our high-risk malware users.

For malware, is there a small percentage of users responsible for a large percentage of events in the same way we saw for phishing? The answer is in Figure 8 below.

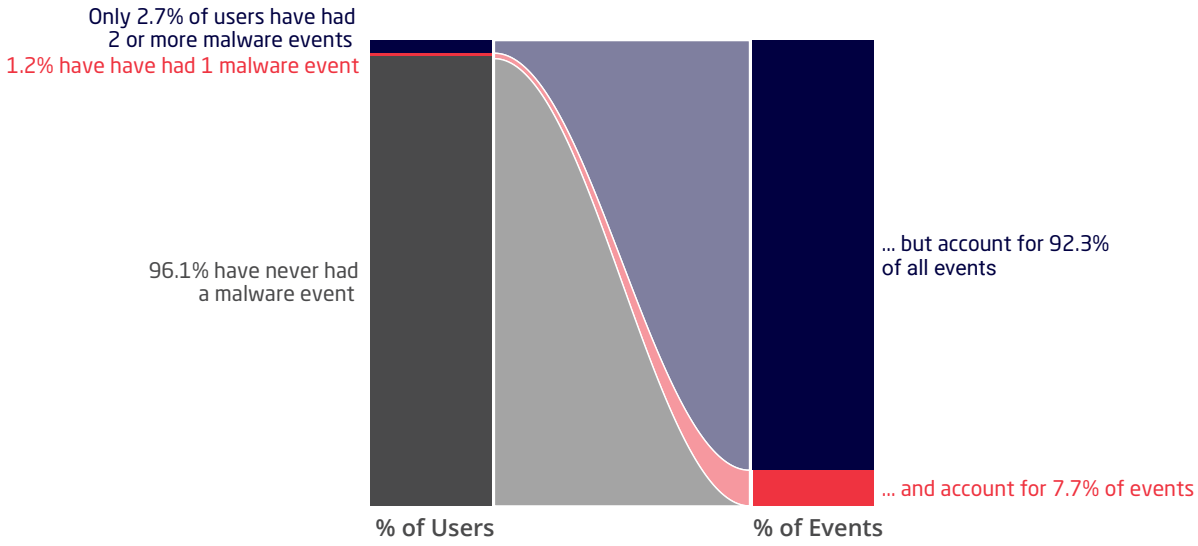

Figure 8: Users responsible for malware events

In Figure 8 we use an "event" to mean the user either downloaded or executed a detected piece of malware. What we can see is an even starker divide than the one we saw with phishing: 92% of malware events are caused by just 3% of users. For malware, the need to focus on just problematic users is even more imperative.

Block rate for malware events

Of course, organizations don't just let malware run amok in their systems. Malware might be nearly as old as the personal computer, but anti-malware controls aren't that far behind. As we said in this section's introduction though, approaches are myriad and different organizations collect data on the effectiveness of those controls in different ways. Some may only record when bad things happen, i.e., execution or download, and don't bother letting us know about blocked malware. Since we're interested in how effective blocking is when it happens, we exclude organizations who don't have any recorded "blocks" of malware.

Without diving too deeply into the data, we can make a few observations. First, there is a great deal of variation on how much malware organizations actually block. In fact, it runs the gambit between everything and nothing, with about 9% of orgs in the former category and 18% in the latter. Perhaps a more interesting nugget here is that even within a single organization malware block rates are not always consistent. Figure 9 shows the department-by-department block rate for three different organizations.

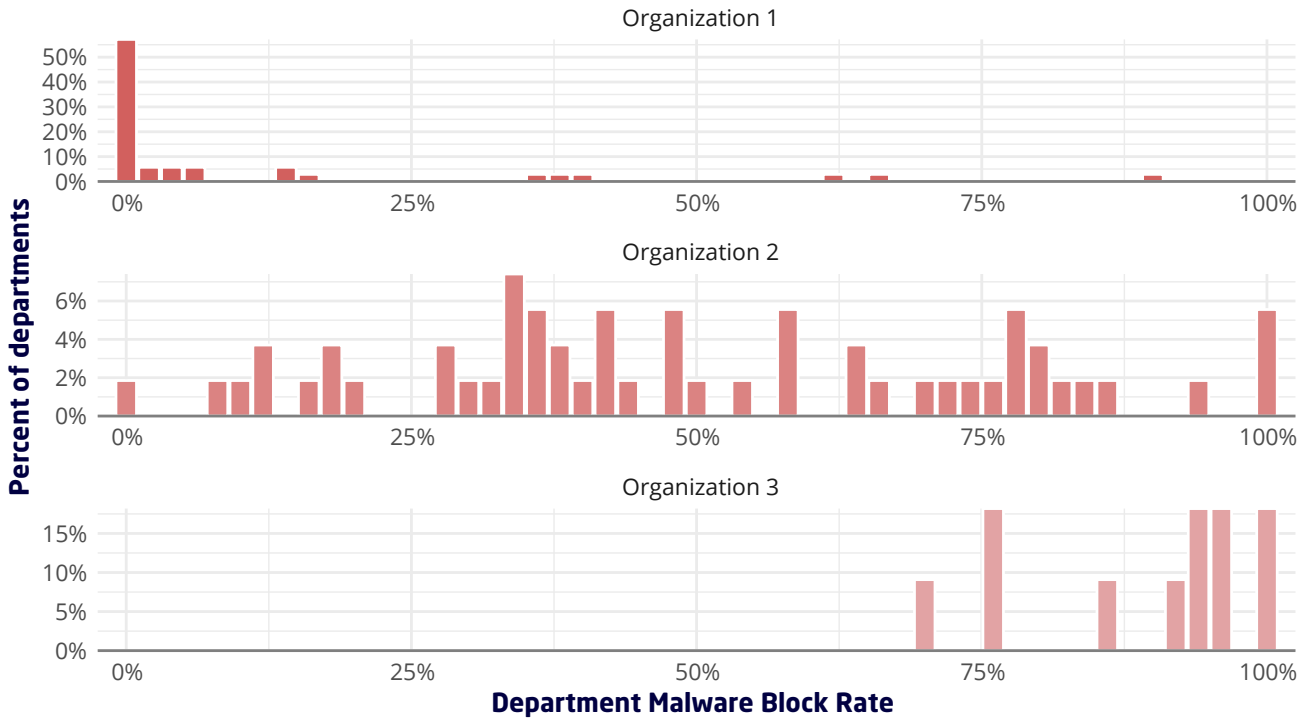

Figure 9: Department block rates for three example organizations

Organization 1 mostly doesn't do any malware blocking, while a few departments are quite successful. Organization 3 is largely successful, with most departments blocking more than 75% of malware. Meanwhile Organization 2 is an absolute grab bag, with departments spanning the range of all possible values from 0% to 100%. Even though organizations may have decided to implement anti-virus across all users, in practice its effectiveness can be highly variable. The major takeaway from looking at these three organizations is that simply making controls available or even requiring them isn't enough. Organizations have to be willing to also measure whether those controls are doing what they are supposed to be doing.

> "
>
> The major takeaway from looking at these three organizations is that simply making controls available or even requiring them isn't enough.

Expected frequency of malware infections

In the same way we approached phishing we can estimate how many malware events departments are likely to have each year. The results can be seen below in Figure 10.

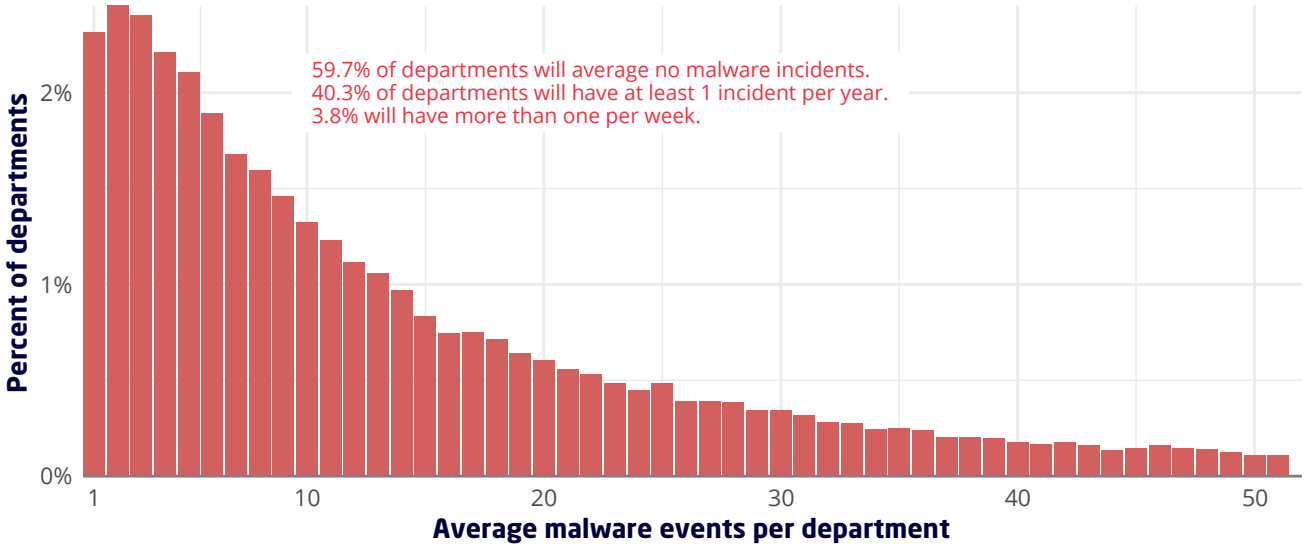

Figure 10: Average malware events per department per year

As before, we see a long tail of danger for some departments, while most are relatively safe. A majority (59%) will have no events because there is no detectable malware, what is detected is blocked before download, or it's never executed after download. Here again we have a long tail though, with 3.6% of departments averaging more than one event per week.

Browsing

TL;DR:

- Secure browsing events probably aren't quite as dangerous as phishing or malware.
- They also happen a lot more.
- A small percentage of users account for most of the secure browsing events.
- Tough to link secure browsing blocks to actual incidents, but it's an indication of risk.

Phishing and malware are both attacker-initiated attacks that can have an immediate effect on an organization's network. But the last type of human risk we want to examine is a bit more of a gray area. The world wide web is dark and full of dangers, even before you get to the parts that people consider the "Dark Web". Questionable websites might merely serve up less than tasteful advertisements or try to direct users to buy not entirely genuine products and services. Not great, but those kinds of things don't pose a direct danger to an organization.

Lurking among those sites are real dangers such as drive-by-downloads and other avenues to phishing sites. Because of the wide variety of possible dangers out there, many organizations enforce secure browsing controls on their network, blocking sites that they deem inappropriate. Now this can include simple time wasters like Facebook to the aforementioned facepalm-inducing attempts to download sketchy files from fraudulent websites. We're not distinguishing amongst the 'badness' here, only trying to get a sense of what a risky user or department looks like when we examine their browsing behavior. With that in mind, let's examine the browsing habits of users and departments.

"

We're not distinguishing amongst the 'badness' here, only trying to get a sense of what a risky user or department looks like when we examine their browsing behavior.

Frequency of browsing violations

Figure 11 displays secure browsing events in a similar fashion to Figure 1 and Figure 7, except this time we do so on a weekly basis rather than a yearly one.

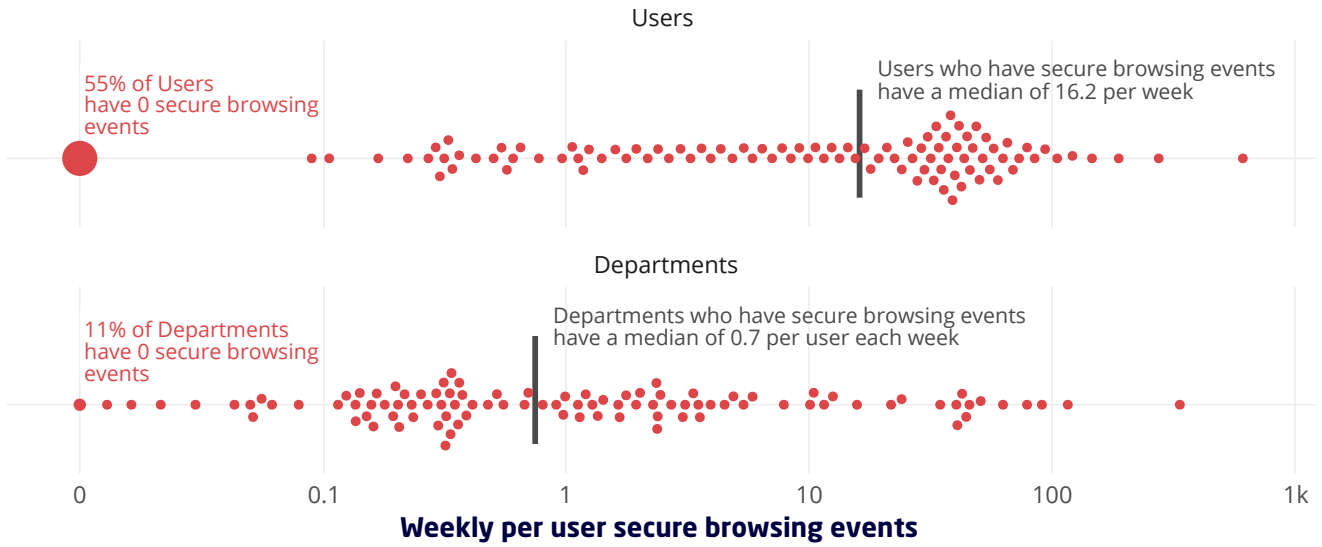

Figure 11: User and Department secure browsing event rates

We again see a divide here among users who don't seem to have any events (about 54%) and those that do. The spread is much greater though, with some users racking up dozens of events per week. In some ways you have to admire the persistence of these users to go back, day in and day out, and attempt to hit blocked websites. Another interesting feature is that at the user level there seem to be two common areas, about once a month and approximately four times a day. This likely reflects different policies, with the lower values coming from orgs focused on "malicious" sites, and the higher values coming from organizations blocking sites that cut into productivity.

Similar to the previous figures, we can see in Figure 12 that a small subset of users is responsible for most of the secure browsing events with just shy of 12% causing 71% of the events. What's interesting here is that a little less than half (45%) have some secure browsing events in their history. This is somewhat less surprising because whom among us, in a fit of boredom, has not spent some time goofing around on the internet at work[7]?

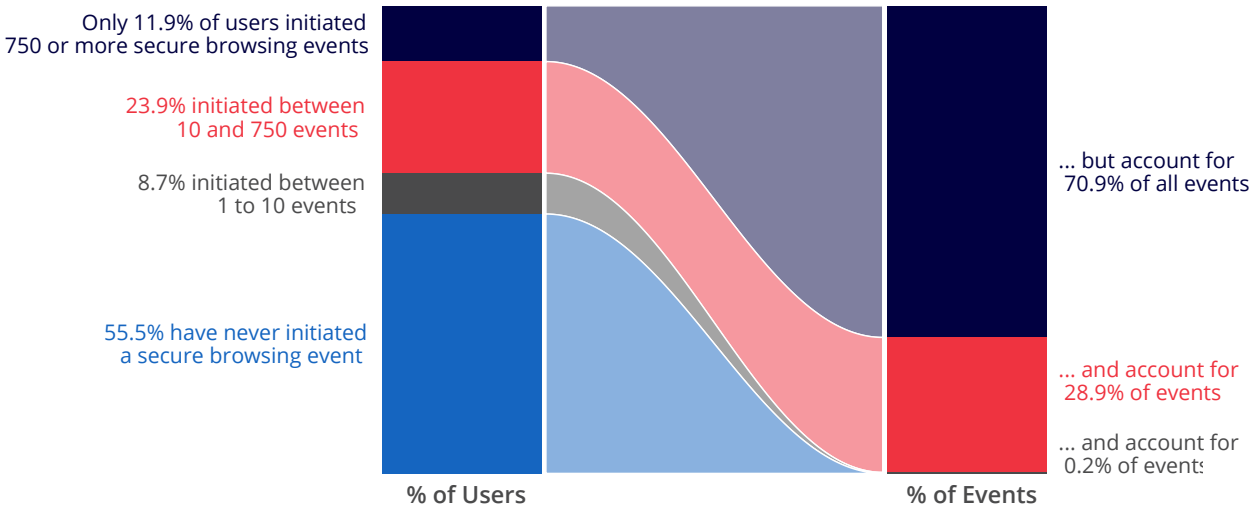

Figure 12: Distribution of secure browsing events per user

Unlike phishing and malware, we don't have much sense of unblocked visits to questionable websites and can't really determine a block rate. Similarly, we don't know when a 'successful' visit to a questionable website results in an adverse event like a malware download, so it's not possible to calculate the same "this is how many bad events per year are the result of insecure browsing".

"

Similar to the previous figures, we can see in Figure 12 that a small subset of users is responsible for most of the secure browsing events with just shy of 12% causing 71% of the events.

7 Boss, if you are reading this, I am definitely part of the 55.5%.

Combinations of Risky Behaviors

TL;DR:

- Risk tends to concentrate in a small percentage of users.
- Users who are 'high risk' in one area are more likely to be high risk in other areas.

Risky phishing users click a few dozen links a year, risky malware users are going to execute about a dozen viruses a year, and those browsing slackers are going to hit tens of thousands of questionable sites a year. But are these all the same users or are the click happy not the same as the ones streaming TV at their desk all day long? After all, if a user was high in all three categories an organization would want to ensure that user got the help and support they needed while applying individualized policies and guardrails for those users to ensure the security of the rest of the network. So, let's take a look.

Identifying the riskiest users

First, we need a good definition of "risky" that can be used across all three categories and is better than "over on the right side of the figure". We define it as follows: A user is "high risk" if they are in the top 75th percentile of users in their organization that have any events. That is, we do not consider those who don't have any events "high risk", but even among those who do, we only look at the top quartile.

We note that this is organization specific in our definition. One organization's 75th percentile might be one malware incident a year, and another's might be 10. But the purpose here is to figure out how many users fall into what combination of risky categories for each organization, lest the different organizational measurement techniques conclude that all the bad users are in the organization that simply measures the most carefully.

"

That is, we do not consider those who don't have any events "high risk", but even among those who do, we only look at the top quartile.

With all that methodology out of the way, based on the definitions above exactly what percentage of users fall into what risky categories? Figure 13 below tells the tale.

- **9% of users are high risk in one category**
- **0.6% of users are high risk in two categories**
- **0.05% of users are high risk in three categories**

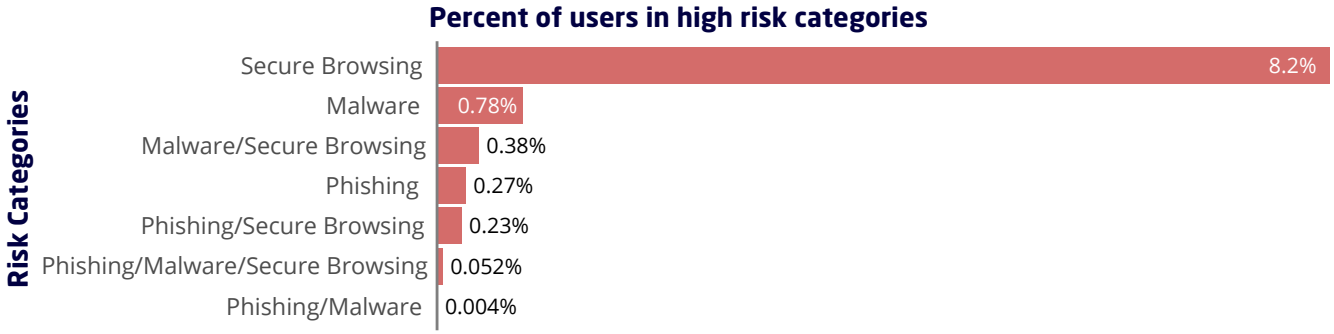

Figure 13: Combinations of risky behavior

Figure 13 certainly shows that risky users tend to be risky across multiple categories of risk. If being risky in, say, malware did not influence secure browsing, we'd expect about 0.2% of users to be in both categories, but approximately 1% are, a nearly five-fold increase. Having users in multiple risk categories seems bad. However, the more concentrated risk is within specific individuals, the fewer organizations have to worry about.

Conclusions

Risk is not uniformly distributed among organizations, that much is clear from the results above. In fact, some users, depending on what exactly we are measuring, represent orders of magnitude more risk than others. So why do we approach human risk as if everyone represents the same risk level? Sending everyone that simulated phishing email or installing AV software everywhere certainly checks the boxes, but it probably doesn't lower your organization's risk profile all that much.

So, you've read this far, seen the pretty figures, and internalized all the numbers, now what? Here are some concrete steps you can take today:

- **Start measuring.** The first step to figuring out which users pose that outsized risk is measurement. This will help establish whether your organization is falling behind in workforce risk or is a superstar.

- **Check the efficacy of your controls.** How many phishing emails are getting through your filters? How uniformly is your AV software installed? Having the controls in place is not enough, you have to make sure they are working properly for everyone.

- **Identify risky users.** Now that you've got all that measurement infrastructure in place, identify who's generating all those security events. This could be users who for some reason are an outsized target for attackers or ones who manage to slip through your security controls or both. And of course, if you find a click-happy user, you might want to check their browsing history.

- **Start helping those risky users.** Because some users represent an outsized risk to an organization, they should also be the focus of an organization's security effort. This may mean setting up guardrails and focused controls.

- **Keep monitoring.** Keep an eye on those risky users and see if their risk profile improves.

This report hopefully gives organizations a yardstick against which to measure themselves. In particular, as they start measuring and checking their controls, they can see if they are on par with other organizations.

As we conclude, you are surely asking, "OK I know what a high-risk user is, but who are they?". Alas reader, you'll have to wait for our next report where we break down which users tend to land in those high-risk categories. Until next time, dear friends.