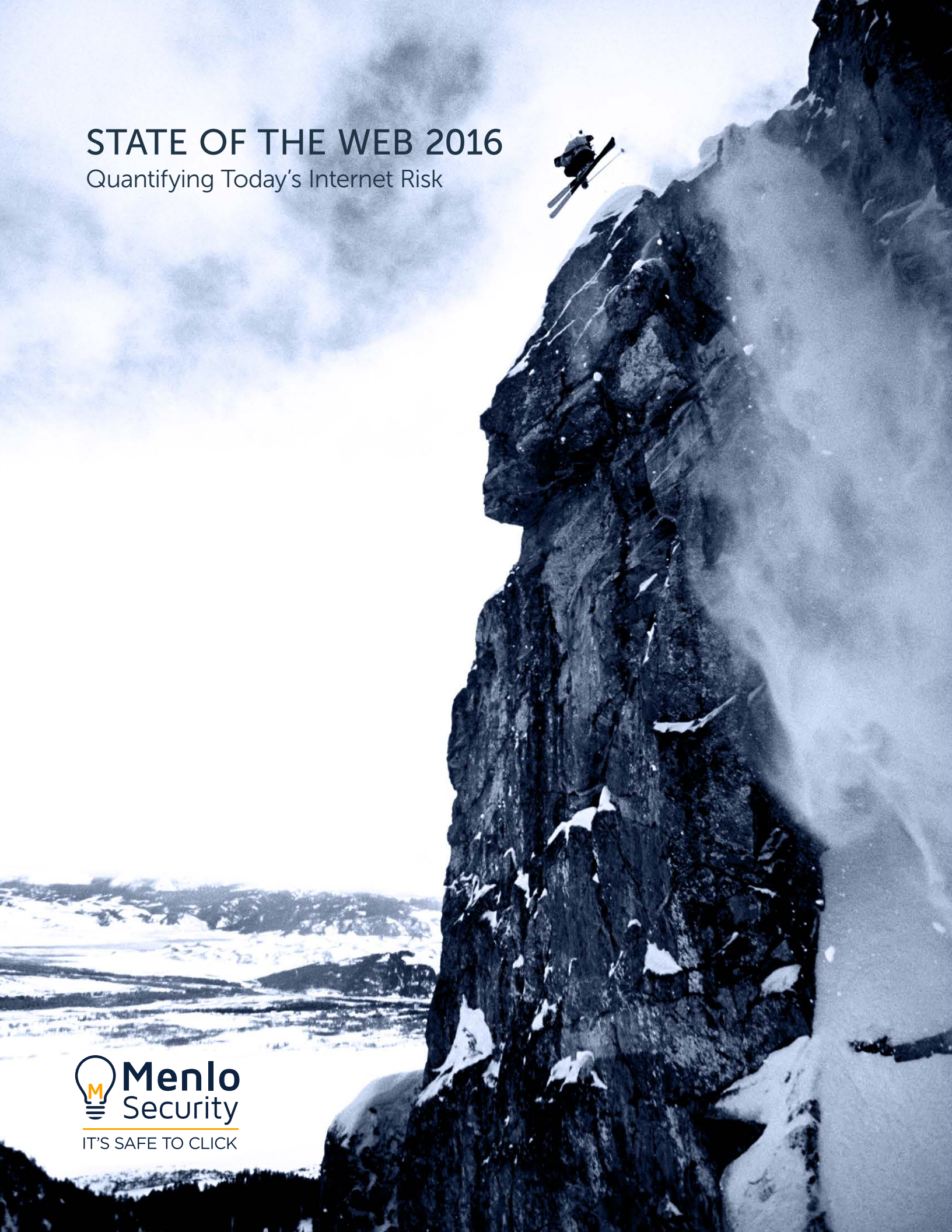
STATE OF THE WEB 2016

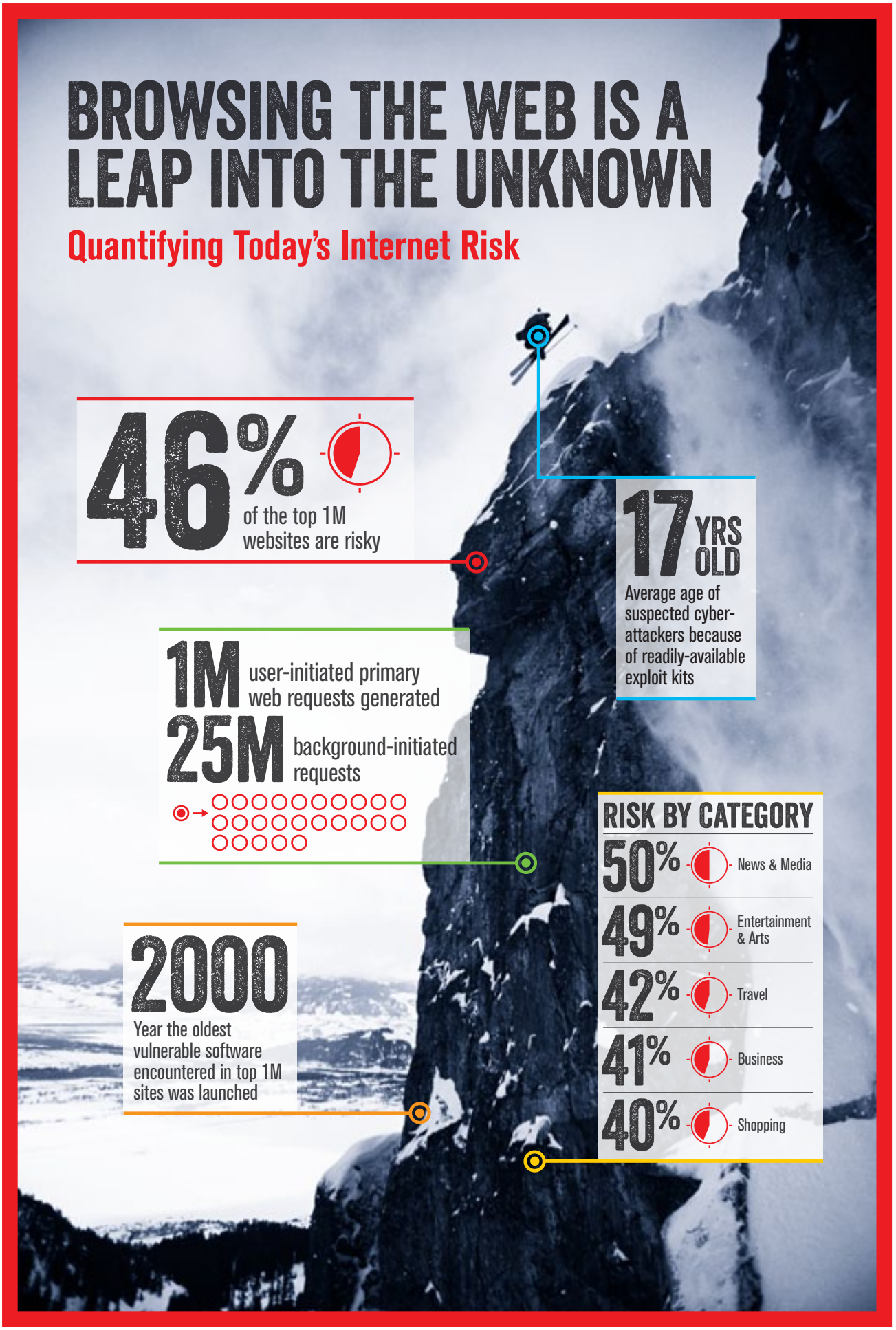
Quantifying Today's Internet Risk



Menlo Security

IT'S SAFE TO CLICK

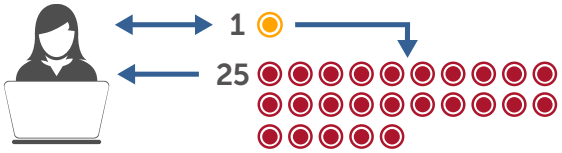
Table of Contents


BROWSING THE WEB IS A LEAP INTO THE UNKNOWN
Quantifying Today's Internet Risk
46% of the top 1M websites are risky
17 YRS OLD Average age of suspected cyber-attackers because of readily-available exploit kits
1M user-initiated primary web requests generated
25M background-initiated requests
2000 Year the oldest vulnerable software encountered in top 1M sites was launched
RISK BY CATEGORY
50% News & Media
49% Entertainment & Arts
42% Travel
41% Business
40% Shopping

Executive Summary

In the last State of the Web report published in 2015, we uncovered two key findings: 1 in 3 domains in the Alexa top 1M are risky, and 1 in 5 domains run vulnerable software. In this report, we again focus on the Alexa top 1 million sites, but also factor in the risks associated with the 25 million requests to background sites that a browser makes when visiting these primary 1M sites. These background sites feed active content to the browser for the purposes of content delivery, trackers, beacons and ad-delivery.

USER VS BACKGROUND-INITIATED REQUESTS

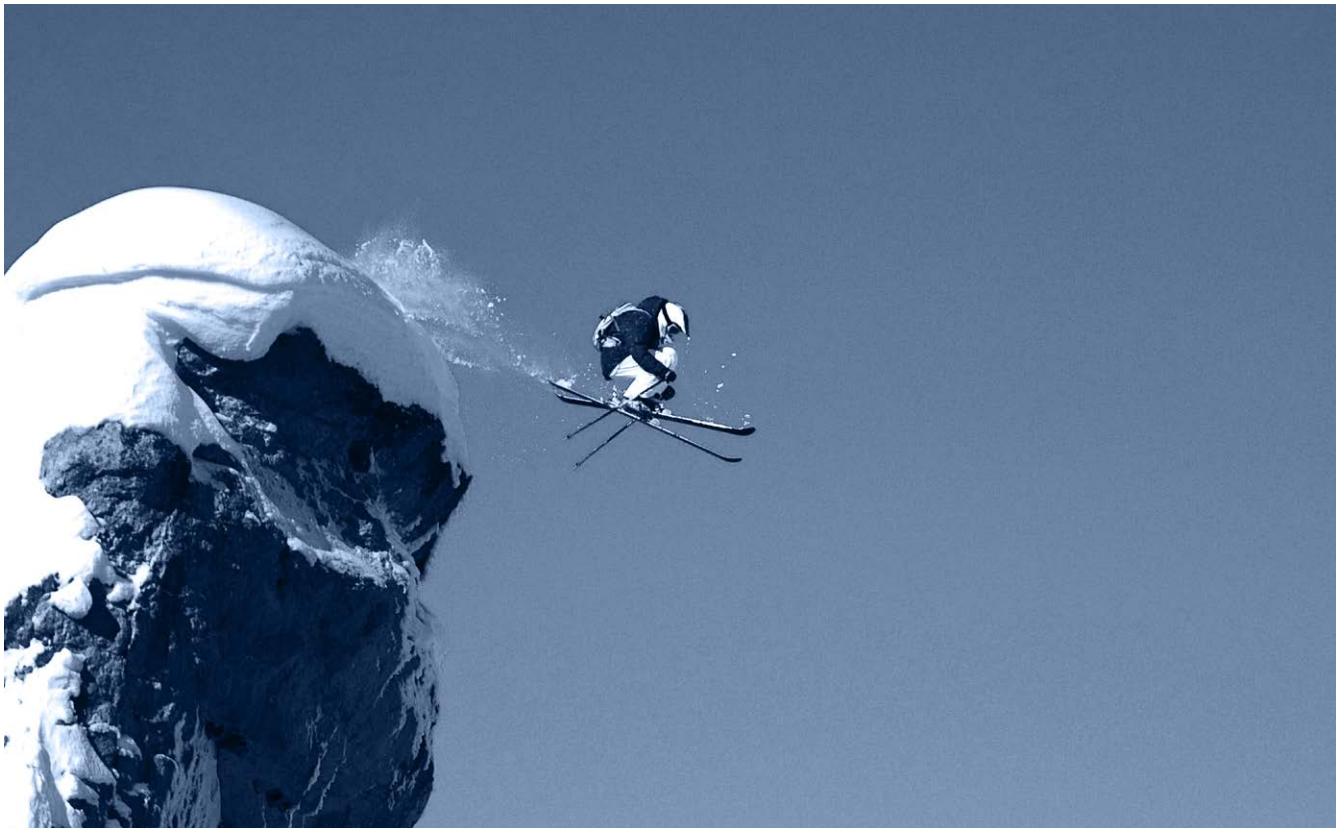


By closely examining key characteristics of the background sites, including software version, release dates, CVE IDs, and third-party risk intelligence, we were able to discern the impact of these background sites on the primary sites' risk.

The results demonstrate that nearly half (46%) of the top million websites are risky. Criminals now have their veritable pick of half the web to exploit. And exploitation is becoming more widespread and effective for three reasons:

1. Risky sites have never been easier to exploit
2. Traditional security products fail to provide adequate protection
3. Phishing attacks can now utilize legitimate sites

Users must recognize that they are taking a significant risk when connecting directly to the Internet. A new approach is needed to address this growing problem.



Introduction

Since the introduction of the modern web browser and JavaScript in the nineties, society has developed an insatiable demand for the media-rich, interactive web experience. Retailers, news outlets, political parties, and more are all competing to gain and hold end user attention to websites. To achieve this end, they present us with vast amounts of information, vivid imagery, streaming video, music, interest-specific advertising, games, etc.

It is increasingly difficult for website owners to host all this content on their own due to scalability, operational management considerations, and the pressure to deliver increasingly rich content. It is for this reason that the vast majority of websites integrate active content from other domains beyond the immediate control of their administrators. We refer to these domains as "background sites." Although this approach does provide the rich interactive experience we expect, it also provides unseen security risks.

One need not look far to see the impact of this risk. In 2016 there were several publicly reported high-profile breaches. AOL and the Huffington Post both served malware. A separate malvertising campaign struck MSN, Telstra, and dating site PlentyofFish.com[1]. Answers.com fell victim to a background site exploit which could have exposed millions of daily visitors to malware. And the New York Times, London Stock Exchange, BBC, Spotify, and The Onion are just a few other notable sites that have been compromised in recent years. As these examples, and this report make clear, there is no such thing as a safe website. The web has become a cesspool of sorts.

Methodology

To analyze the top 1 million web sites, Menlo Security developed a distributed Chrome-based browser farm to load the homepage of each of the web sites. The Chrome browser was further instrumented with a proprietary Menlo Security Risk Analyzer extension that could monitor the loading and execution of JavaScript. Using a real browser to load the webpage was critical to the page loading in the same manner it would for an end user. This also triggered software execution in the form of scripts as well as dynamic loading of ads, beacons, trackers, etc. In addition to tallying up the number of scripts in each page, the Menlo Security Risk Analyzer was also able to passively fingerprint the script origin servers (website software), categorize these background sites, as well as correlate the CVE-IDs of these sites based on the website software. We also leveraged readily available threat intelligence feeds to determine whether sites were categorized as a "known-bad," or had security incidents in the last 12 months.

Based on this information, Menlo Security classified sites as risky if one or more of the following was true:

- **Homepage or background site was running software with known vulnerabilities (CVEs)**
- **Homepage or background site was categorized as "known-bad" such as phishing, malware sites, etc.**
- **Homepage or background site has had a security incident in the last 12 months**

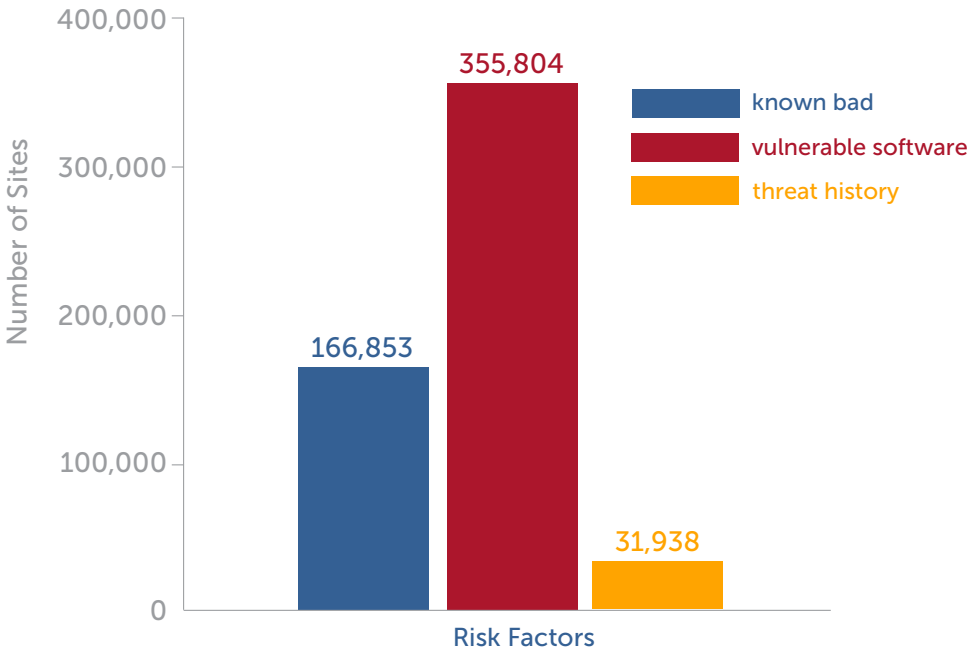
[1]2016 Verizon DBIR

Risk Factors

Menlo Security considers a site risky if either the homepage, or associated background sites, is running vulnerable software, is known-bad, or has had a security incident in the last 12 months. By a factor of more than 2, vulnerable software was the leading factor in classifying a site as risky. Of the 1 million sites, 355,804 were either running vulnerable software or accessing background domains running vulnerable software; 166,853 fell into known-bad categories, while 31,938 experienced a recent security incident.

Note: Some vulnerable sites fall into multiple risk categories.

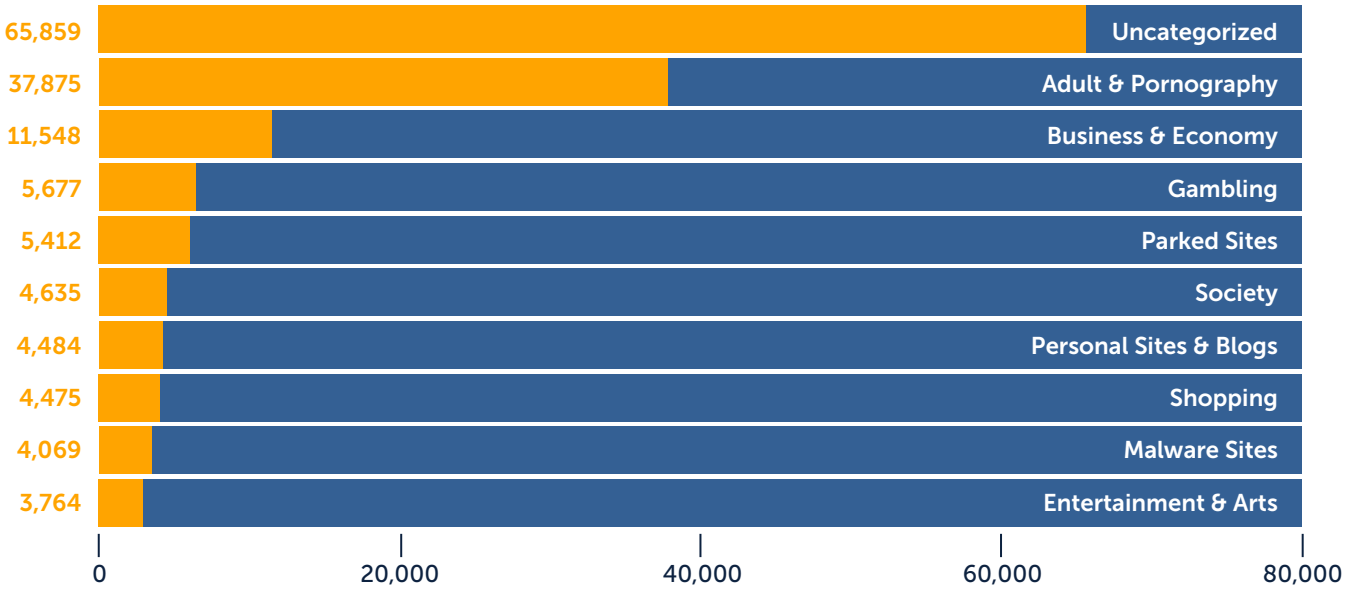
BREAKDOWN OF RISK



Top Known-Bad Categories

Although some of the most popular categories of known-bad sites would be expected, such as pornography and gambling, others, such as news and shopping, are generally trusted by multitudes of businesses and the general public. Of particular interest is the fact that uncategorized sites, those that are unknown to URL categorization services, are the leading source of known-bad. This underscores the challenge that enterprises face when enforcing policies for such sites. By allowing access to uncategorized sites, businesses put themselves at unnecessary risk, incur high costs to sanitize infected machines, and waste money chasing false positive security alerts. If they deny access to uncategorized sites, administrators are swamped with re-categorization requests, which is a manual process that often requires hiring more staff to cope with policy updates, URL re-categorisation requests, and keeping impacted staff updated on all tickets.

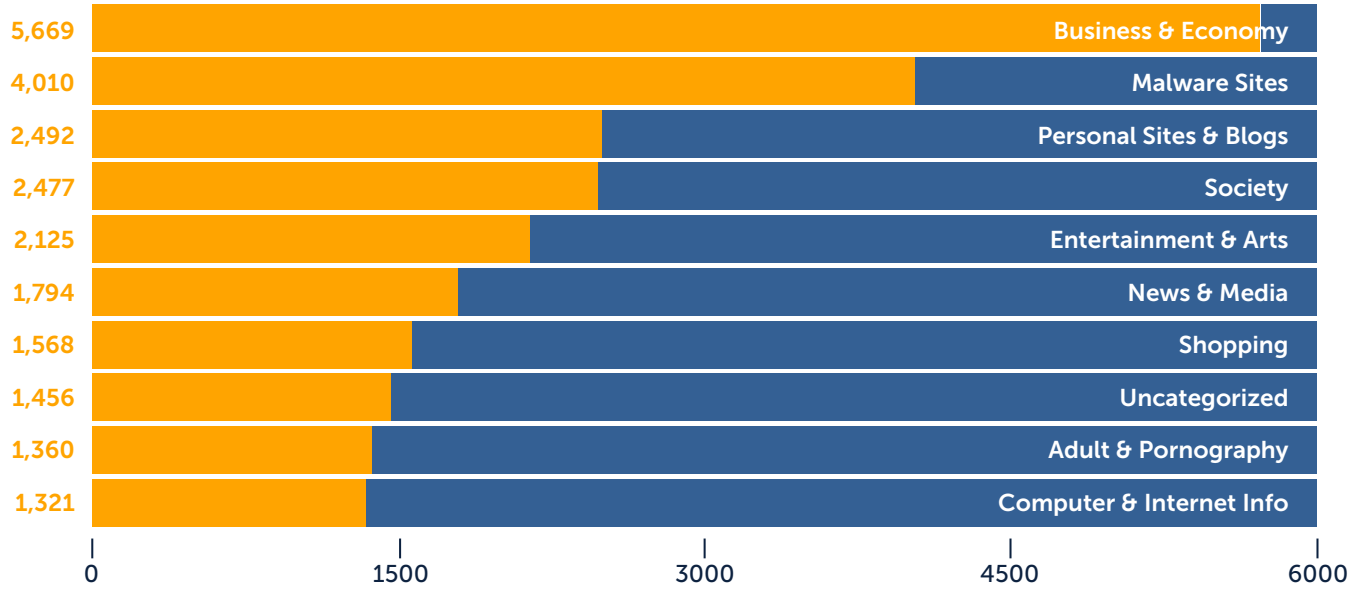
TOP CATEGORIES OF KNOWN-BAD SITES



Categories with Recent Security Incidents

Again we see categories we would expect to have recent security incidents, such as pornography and malware sites. But the vast majority of recent incident categories are ones that an average person would visit while at work, as part of their daily routine. Whom amongst us doesn't check the news and weather each morning? Or get the latest updates on the rich and famous? Or catch up on our shopping, read our favorite blogs, or watch a viral video? Risk is ever-present, even with the most trusted "legitimate" sites.

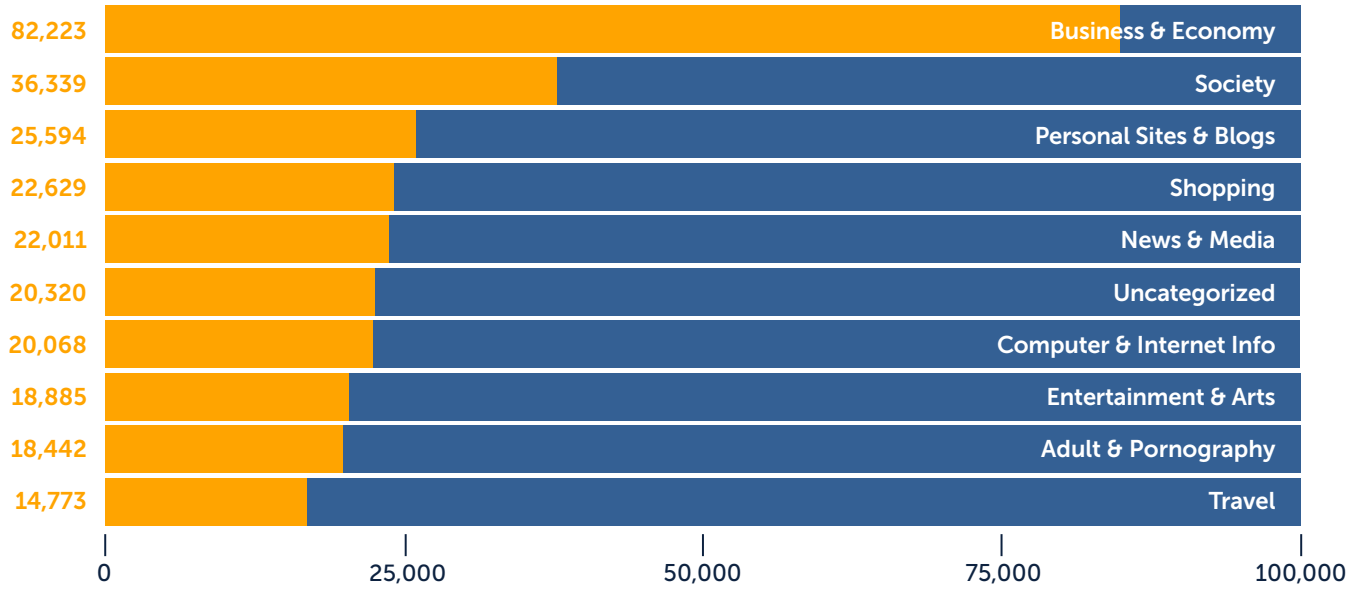
TOP CATEGORIES OF SITES WITH RECENT THREAT HISTORY



Categories with Vulnerable Software

Looking at the top 10 categories in which vulnerable software is most prevalent, Business & Economy sites are the uncontested winner, with more than 3 times as many vulnerable sites as the Adult & Pornography category.

TOP CATEGORIES OF SITES WITH VULNERABLE SOFTWARE

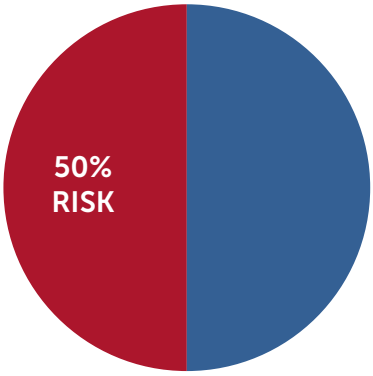


Surprisingly Risky Categories

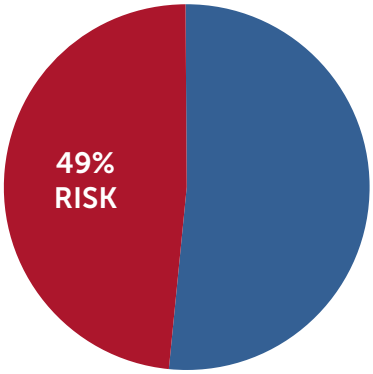
With an understanding of the most vulnerable categories, we can calculate a category's risk as the ratio of vulnerable sites to total sites. The top three riskiest categories are News & Media, where 50% of sites satisfy at least one of our three criteria, followed by Entertainment & Arts at 49%, and Travel at 43%. The least risky category of the top 10, Computer & Internet Info, still comes in at a massive 37%.

TOP RISK CATEGORIES

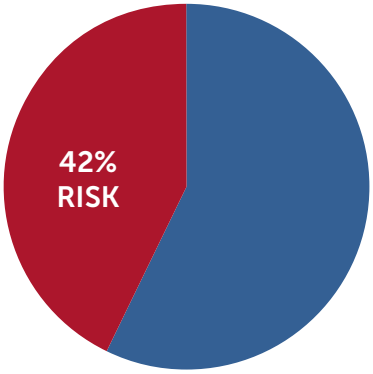
NEWS & MEDIA



ENTERTAINMENT & ARTS



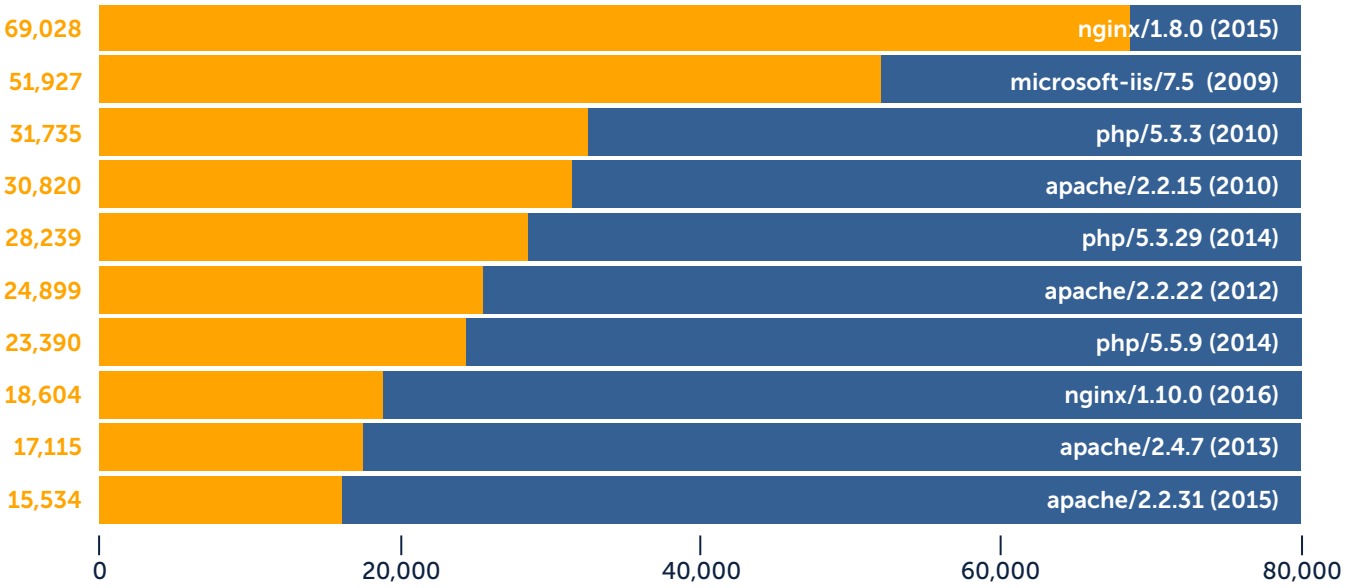
TRAVEL



Most Prevalent Vulnerable Software

All software has vulnerabilities, and the vendors listed below are all diligently providing patches to close vulnerabilities as they are discovered. The key is for site owners to be equally diligent about applying those patches. Old versions of software pose a material risk.

MOST USED VULNERABLE SOFTWARE



Implications

What does all this mean? If nearly 1 out of 2 sites is risky, why are you not getting infected every day? The fact is there are currently more vulnerable websites than attackers to exploit them. Additionally, cyber criminals run rotating campaigns in shifting locations to avoid detection. The dynamic nature of these campaigns coupled with the ratio of attackers to sites makes it nearly impossible for every site to be infected at all times. What is important to understand, however, is given the current state of the web, villains have their veritable pick of half the web to exploit. And exploitation is becoming more widespread and effective for three reasons:

1. Risky sites have never been easier to exploit
2. Traditional security products fail to provide adequate protection
3. Phishing attacks can now utilize legitimate sites

Risky Sites Have Never Been Easier to Exploit

Gone are the days when cyber criminals required a great deal of technical sophistication to launch an attack against a network or endpoint. Today, exploit kits are readily available to anyone, as are the instructional videos that provide step-by-step execution instructions. The expertise requirement has all but vanished. Underscoring this point, the average age of suspected cyber-attackers dropped from 24 to 17 over the course of 2015.[2]

The combination of wide spread software vulnerabilities, pervasive exploit kits, and throngs of new attackers has created the perfect storm. For example, Microsoft-iis 7.5 was the second most common vulnerable software seen in our report, and is currently running on over 50,000 of the top one million websites. For an individual to compromise a web server running this software, it is a simple matter of using exploit kits readily available on the Internet to enable a total system compromise. Today, within minutes, any motivated attacker can exert full control over a primary or background site, and deliver ransomware to unsuspecting visitors.

Traditional Security Solutions Fail to Provide Adequate Protection

Exacerbating the risk issue, the vast majority of malware prevention products ignore background sites. They attempt to prevent attacks by distinguishing between "good" and "bad" elements, and then implement policies intended to allow "good" content and block the "bad." The basic approach is the same regardless of the detection method being used: From simple anti-virus signatures and web filtering, to virtual execution and sandboxing, and even newer approaches like big data analysis and behavioral modelling, all ultimately come down to an "allow" or "block" decision. In every case, the detection is never perfect, and thus the policy choice involves a level of risk that the wrong decision is being made.

No technology makes the right "good" vs. "bad" decision 100% of the time, and this leads to mistakes that can be very costly. False positives cause alerts that need to be analyzed and also result in calls to the help desk to unblock legitimate content. False negatives enable malware or malicious emails to get through. Once an endpoint is compromised, it can be controlled by attackers to enable access to intellectual property, steal information, act in a botnet, and many other criminal activities. End users are often fooled into clicking on links in phishing emails that take them to sites where they inadvertently share account names, passwords, and other sensitive data. These mistakes are costly and occurring at an alarming rate.

[2] https://www.theguardian.com/technology/2015/dec/08/average-age-of-cyber-attack-suspects-drops-to-17

Phishing Attacks Can Now Utilize Legitimate Sites

The most common element of recent high-profile breaches is phishing. Although traditional phishing attacks involve the creation of a new imposter, or "spoofed," site, the sheer volume of vulnerable trusted sites makes it very easy for attackers to compromise a legitimate site and send that link as part of a phishing campaign. With this approach, attackers no longer need to worry that URL filtering will thwart their efforts, and they avoid anomalies in the link address, such as misspellings, special characters, or numbers that might raise suspicions. Clicking on what is a perfectly legitimate link within such a phishing email can expose a user to a drive-by malware exploit that could deliver ransomware, or mark the beginning of a larger breach.

Global Response

The current state of the Internet risk is eliciting responses from businesses and governments around the world. In one of the more extreme examples, Singapore's government announced in June of 2016 that it would cut off Internet access for government workstations within a year for security reasons, a surprise move in one of the world's most wired countries.[3] Fortunately, options exist today that are far less impactful to users and productivity, but equally effective in reducing risk.

[3]http://www.securityweek.com/singapore-blocking-internet-access-government-computers

Recommendations

Based on the information in this study we provide recommendations to three audiences: enterprise IT administrators, website owners, and end users.

Enterprise IT Administrators

Traditional IT security products, such as secure web gateways, are blind to the majority of risk summarized in this report. New threat prevention techniques, such as isolation and remote browsing, are advocated by security analysts. Isolation inserts a secure, trusted execution environment, or isolation platform, between the user and potential sources of attacks. By executing sessions away from the endpoint and delivering only safe rendering information to devices, users are protected from malware and malicious activity regardless of the risk-level of any site. Menlo Security recommends the addition of threat isolation technology to every security architecture to eliminate the risk associated with all sites, including background sites. We are not alone in this recommendation. Gartner recommends isolation as part of its Adaptive Security Architecture, and highlights isolation in the *It's Time to Isolate Your Users From the Internet Cesspool With Remote Browsing* report.

Website Owners

For website owners concerned with potential risk, the first and most obvious step is to be sure servers are running the latest updated software versions that have mitigated known CVEs. To mitigate risk on background domains beyond the website owner's immediate control, we recommend implementing measures such as Content-Security-Policy (CSP) which prevents the execution of malicious content on a trusted website. Other new browser security mechanisms complementary to CSP include:

- **Patch Vulnerable Software** and run the latest versions that have mitigated known CVEs
- **Sub-Resource Integrity (SRI)**, to ensure only known, trusted resource files (typically JavaScript, CSS) are loaded from third-party servers (typically CDNs)
- **Mixed Content**, to clarify the intended browser's policy on pages loaded over HTTPS and linking content over plaintext HTTP
- **Upgrade Insecure Requests**, hinting browsers on how to handle legacy links on pages migrated to HTTPS
- **Credential Management**, a unified JavaScript API to access user's credentials to facilitate complex login schemes

End-Users

For end-users, Menlo Security recommends the following best practices:

- **Look closely at the URL** – special characters, numbers, etc. In many cases what looks like Paypal may not actually be a Paypal site. If your bank or a site that you have an account with sends you a 'password reset' email, login to the bank directly instead of clicking on the link in the email
- **Use an ad-blocker if you can** – Malvertising campaigns are on the rise and the ad sites typically are infiltrated first to deliver Ransomware to unsuspecting users
- **Use the Chrome browser** – Amongst browsers, Google Chrome has a higher degree of focus on security, and more importantly, contains an implementation of Flash that is highly restrictive
- **Disable or uninstall Flash** – Flash is known to carry dangerous malware payloads. Users don't often need Flash and most sites have switched to HTML5 video
- **Keep the software on your PC, Mac, and smart phone up to date** – Companies frequently update software for security vulnerabilities
- **Do not download PDFs or Word documents from untrusted sources** – Ransomware predominantly spreads via weaponized documents
- **Google offers a free PDF URL converter** that can be used to safely view the contents of a PDF without having to download it
- **Do not download executable and Zip files from untrusted sources** – There's no way to guarantee the efficacy of these files and if you go ahead and download it anyways, you deserve to pay bitcoins
- **Use browser-based web email instead of an email client** –The main advantage is most web mail providers have document previews (like Google Docs) that can be used to safely view the document without having to download it first
- **Avoid custom apps and extensions, especially from untrusted sources** – In the recent past, many extensions and apps that have started out good have gone rogue as part of a software update. Do not install apps from untrusted stores or sources as you have no way of verifying if they have been weaponized

Menlo Security
IT'S SAFE TO CLICK

934 Santa Cruz Avenue
Menlo Park, CA 94025
Tel: 650 614 1795
info@menlosecurity.com

menlosecurity.com