SEARCHLIGHT CYBER

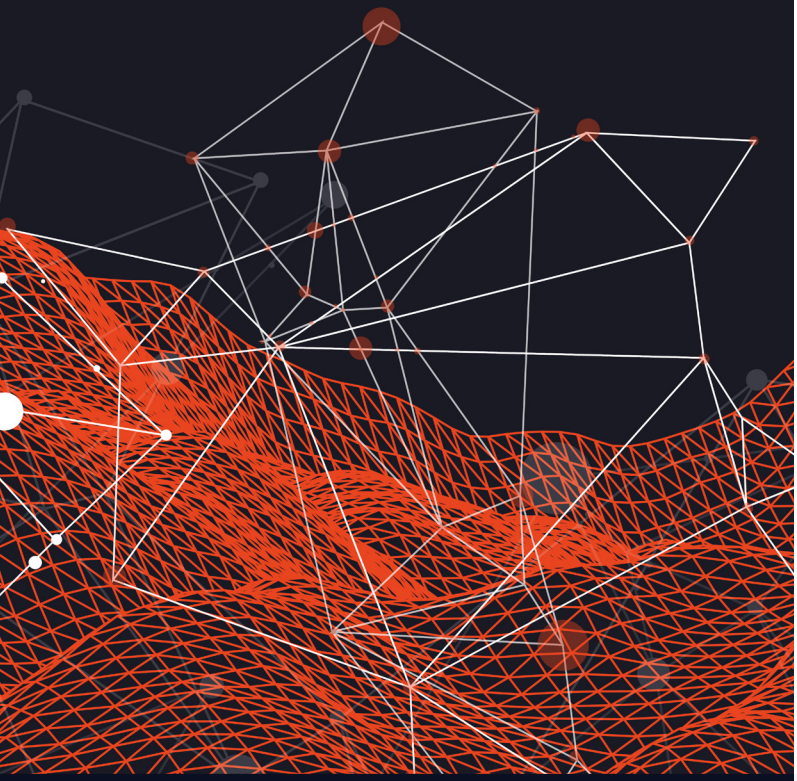
THE CORRELATION BETWEEN DARK WEB EXPOSURE AND CYBERSECURITY RISK

ISO 27001 INFORMATION SECURITY MANAGEMENT SYSTEM

CYBER ESSENTIALS

AICPA SOC

Crown Commercial Service Supplier



SEARCHLIGHT.
CYBER

Searchlight Cyber provides organizations with relevant and actionable dark web intelligence, to help them identify and prevent criminal activity. Founded in 2017 with a mission to stop criminals acting with impunity on the dark web, we have been involved in some of the world's largest dark web investigations and have the most comprehensive dataset based on proprietary techniques and ground-breaking academic research. Today we help government and law enforcement, enterprises, and managed security services providers around the world to illuminate deep and dark web threats and prevent attacks.

ISO 27001 INFORMATION SECURITY MANAGEMENT SYSTEM
CYBER ESSENTIALS
AICPA SOC
Crown Commercial Service Supplier

ABOUT THE MARSH MCLENNAN CYBER RISK INTELLIGENCE CENTER

The Marsh McLennan Cyber Risk Intelligence Center (Center) is Marsh McLennan's enterprise-wide cyber data, analytics, and modelling center of excellence. The Center was founded in 2021 with a mission to advance how businesses and their communities quantitatively and economically anticipate, measure, and manage cyber risk. By leveraging advanced analytical and modeling techniques, the Center brings together Marsh McLennan's expansive proprietary data and models across its Marsh, Guy Carpenter, and Oliver Wyman businesses with complementary leading external sources, to develop a robust suite of cyber quantification tools. The Center's tools power cyber modeling exercises, cyber analytics, and thought leadership insights for Marsh McLennan clients around the world, including cybersecurity technology organizations, insurance and reinsurance providers, and others.

CONTENTS

EXECUTIVE SUMMARY

The presence of any data relating to your organization on the dark web demonstrably increases your risk of a cyberattack.

That is the core finding from the Marsh McLennan Cyber Risk Intelligence Center's comprehensive study, which analyzed our dark web dataset against a sample of more than 9,000 organizations.

Of that sample, 3.7 percent of the organizations had suffered one or more cyber insurance losses in the last four years. What the Marsh McLennan team wanted to determine was whether those breaches had a higher likelihood of occurring, based on dark web intelligence in the year leading up to the incident.

PRE-ATTACK INTELLIGENCE

The dark web is an obfuscated part of the internet that is prolifically used by cybercriminals to communicate between one another, plan their attacks, and buy, sell, and build the tools they need to execute them.

This activity is known as the "pre-attack" phase of a cybersecurity incident: the actions that cybercriminals undertake before they launch their campaign against an organization and breach their network.

It stands to reason that the presence of this pre-attack activity against a specific organization would mean that they have an increased likelihood of being the victim of a cybersecurity incident - and that is exactly what Marsh McLennan's study confirmed.

In this report, Marsh McLennan's team demonstrated a statistically significant correlation between all of our dark web intelligence sources - including (but not limited to) dark web market listings, hacking forum chatter, and dark web traffic to and from the corporate network - and an increased likelihood of suffering a cybersecurity incident. Put simply: the presence of any dark web findings related to an organization - without exception - was associated with a higher likelihood of a breach.

While every single source shows an increased risk of a cyberattack, organizations need visibility of all the sources to truly understand their cybersecurity risk. Soley gathering intelligence on compromised user accounts, for example, does not provide the whole picture if you can't see whether the organization is listed on a dark web marketplace. This study shows through a multi-variable analysis how multiple dark web intelligence sources can lead to a more reliable estimate of combined cybersecurity risk.

ACTING ON CYBERSECURITY RISK

The analysis confirms that dark web intelligence is highly correlated with forthcoming cyber incidents. What matters now is how organizations act on this information.

The first step has to be to gain visibility into your exposure on the dark web. Understanding where the organization is vulnerable is critical for informing defense and the value of pre-attack intelligence is that it creates an invaluable window of time for the security team to act before the network is breached. If the exposure is identified early enough, the company can take action to prevent the cybersecurity incident.

Furthermore, all security teams have pressure on resources, which means they have to decide where to prioritize their efforts based on the highest levels of risk. Marsh McLennan's study has determined which dark web intelligence sources are most correlated to cybersecurity incident frequency, which can help organizations to determine the dark web threats they should focus on.

Once visibility into threats emerging from dark webs is established, it is then critical that this exposure is continuously monitored. The dark web is anything but static; new sites emerge every day, thousands of posts are written on hacking forums, new products are bought and sold on illicit markets. An organization's dark web exposure will fluctuate over time and identifying new threats quickly is the key to mitigating risk and reducing the chances of a cybersecurity incident.

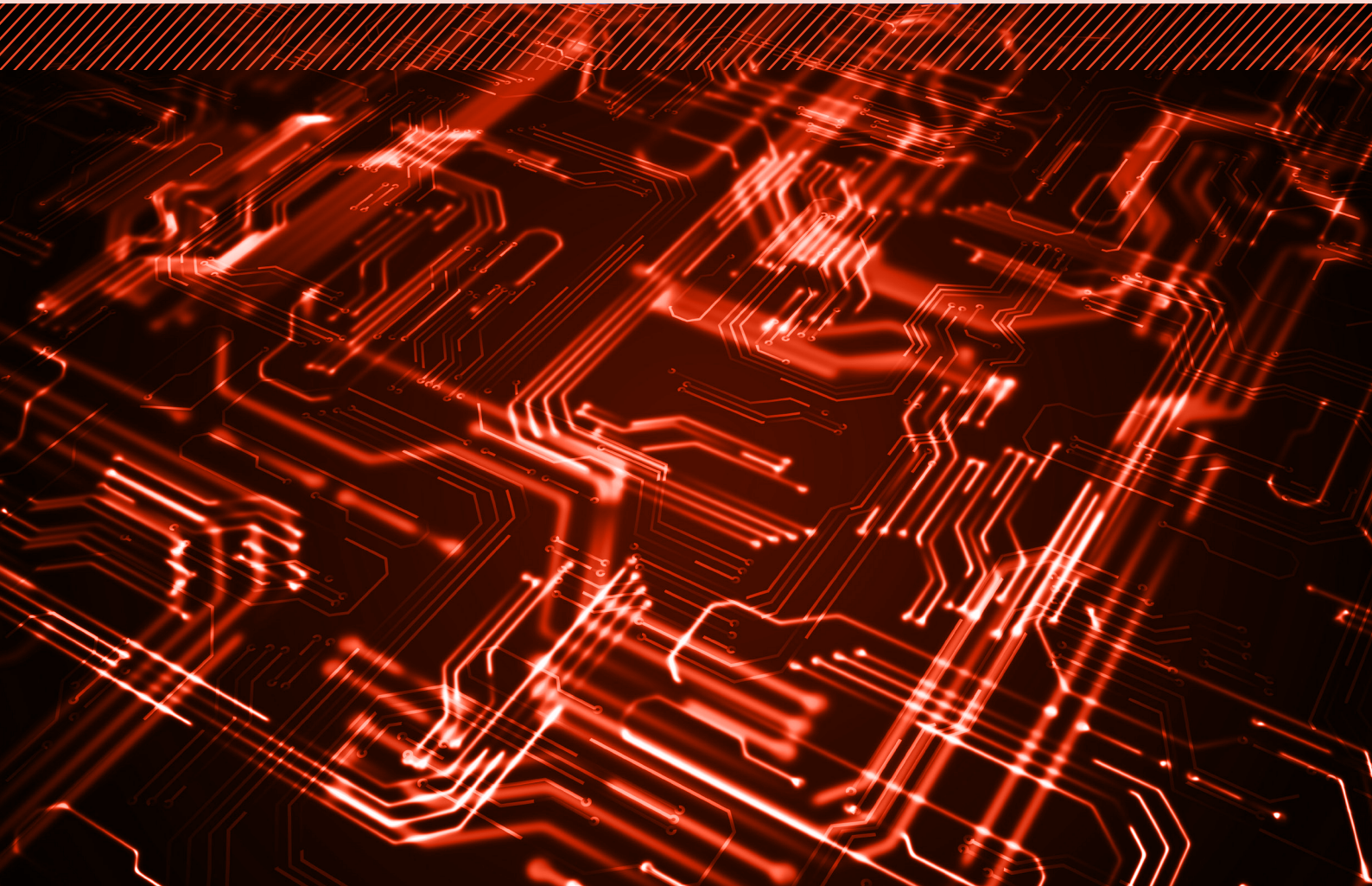
BEN JONES
Co-Founder and CEO
Searchlight Cyber

INTRODUCTION

Marsh McLennan's Cyber Risk Intelligence Center conducted this study to determine if intelligence sources from the dark web were correlated with the frequency of cyber insurance claims. Searchlight Cyber was able to generate hundreds of results related to the organizations Marsh McLennan assessed, and showed a statistically significant correlation to cyber incident data.

SCOTT STRANSKY
MANAGING DIRECTOR AND HEAD OF THE CYBER RISK INTELLIGENCE CENTER, **MARSH MCLENNAN**

> Historically the insurance industry has focused on data from within an organization, such as questionnaires, along with outside-in technographic scans for determining cybersecurity risk. While this data is extremely valuable, ignoring dark web factors external to the organization's network leaves the industry with a blind spot around who could be targeting the organizations they insure and the resources those cybercriminals possess to execute their attacks.
>
> Our analysis of the dark web intelligence market found that this dataset is highly correlated with cyber insurance loss frequency. We are delighted to be able to share the results of the analysis we conducted with Searchlight Cyber, which showed that external threat factors are correlated with cybersecurity incident frequency.



METHODOLOGY

Marsh McLennan aimed to establish the correlation between data breaches and findings on the dark web for organizations a year before the breach, as collected by Searchlight Cyber. The study involved 9,410 organizations with an overall breach rate of 3.7 percent from 2020 to 2023. It consists of two major analyses:

1. **Examining the impact of findings in dark web intelligence sources one-by-one (single-variable analysis)**
2. **Examining the impact of findings in dark web intelligence sources jointly (multi-variable analysis)**

KEY FINDINGS

ALL NINE OF SEARCHLIGHT CYBER'S DARK WEB INTELLIGENCE SOURCES WERE FOUND TO HAVE A **STATISTICALLY SIGNIFICANT CORRELATION** TO INCREASED CYBERSECURITY RISK.

THE INDIVIDUAL FACTORS THAT WERE CORRELATED WITH THE LARGEST INCREASE IN THE CHANCE OF A CYBER INSURANCE CLAIM WERE **COMPROMISED USERS (2.56X)**, **DARK WEB MARKET LISTINGS (2.41X)**, AND **OUTGOING DARK WEB TRAFFIC (2.11X)**, RELATIVE TO ORGANIZATIONS WITH ZERO FINDINGS IN THOSE FACTORS.

COMBINING MULTIPLE DARK WEB SOURCES PROVIDES A STRONGER INDICATION OF INCREASED CYBER RISK THAN IN ISOLATION.

PASTE RESULTS, **OSINT RESULTS**, AND **DARK WEB MARKET LISTINGS** WERE FOUND TO BE THE MOST CORRELATED TO INSURANCE CYBER LOSS FREQUENCY **IN CONJUNCTION WITH OTHER FACTORS**.

DARK WEB INTELLIGENCE SOURCES EXPLAINED

Marsh McLennan's analysis found a significant statistical correlation between all of Searchlight Cyber's dark web intelligence sources and the increased risk of a cybersecurity incident within the next 12 months.

DARK WEB MARKET LISTINGS

The mention of the organization or data related to the organization on a dark web market. Dark web markets operate in much the same way as regular online marketplaces such as Amazon or eBay but specialize in illicit goods that cannot be sold on the clear web. Cybercriminals use dark web marketplaces to sell company data, as well as access to systems and infrastructure.

FORUM POSTS

The mention of the organization or data related to the organization on a dark web forum. Forums are often used by threat actors to discuss tactics, techniques, and to share exploits when targeting an organization. For example, Initial Access Brokers - a type of cybercriminal that specializes in selling access to an organization's network - often hold auctions on dark web forums.

COMPROMISED USERS

Compromised accounts on the dark web related to an organization. Data such as passwords, email addresses, and usernames can be obtained through a number of means, including through phishing campaigns, the use of infostealer malware, or through a third party data breach if employees have used work credentials for other services. Stolen username and password combinations are one of the most common vectors of a cyberattack.

TELEGRAM CHATS

The mention of the organization or data related to the organization on Telegram, a communication platform that is commonly used by cybercriminals to conduct pre-attack activity. Telegram channels are used to share cybercriminal and fraud techniques, sell hacking tools, and leak stolen databases.

INCOMING DARK WEB TRAFFIC

Traffic originating from the dark web and connecting to an organization's infrastructure. This means that someone from the dark web is visiting the network. Depending on the component they are connecting to and the size of the data, this could be an indicator of malicious activities ranging from cybercriminal reconnaissance, to port scanning, all the way up to malware installation.

OUTGOING DARK WEB TRAFFIC

Traffic originating from the organization's network and connecting to the dark web. This means that someone within the corporate environment is calling out to the dark web. This could be an indicator of malicious activity taking place on the network, such as Command and Control (C2) malware beaconing out, or a malicious insider visiting dark web forums, marketplaces, or sites.

DARK WEB PAGES

The mention of an organization or data related to an organization on a dark web site. Dark web pages are similar to websites on the clear web but are hosted through networks such as The Onion Router (Tor) to provide the site creator and its visitors with anonymity. As such, dark web pages typically contain a wide variety of illicit or illegal activity.

PASTE RESULTS

The mention of an organization or data related to an organization on plain-text repositories that are designed to facilitate the sharing of large blocks of computer data in online forums. Cybercriminals are known to post stolen or illegal information on paste sites, including organizational data that has been captured in data breaches.

OSINT RESULTS

Open Source Intelligence (OSINT) results are extracted from all Searchlight Cyber data feeds and represent all of the assets related to an organization - such as IP addresses and web domains - that have been identified on the dark web. These assets could represent a point of vulnerability that cybercriminals are looking to exploit.

HOW EACH DARK WEB SOURCE INDIVIDUALLY IMPACTS CYBER RISK

Marsh McLennan examined the impact of findings in each of the dark web fields individually in a single-variable analysis, demonstrating that a result in any field in the previous 12 months increased the risk of a cyber insurance loss. This statistical correlation was established using confidence intervals (see methodology in **Appendix 1**).

Figure 1 shows the percentage of organizations who had a breach with and without a finding in each category, and clearly demonstrates the impact each dark web finding has on the level of risk. It shows that organizations who appeared in Dark Web Market Listings had the highest breach rate (8.67 percent), followed by those with presence in Telegram Chats (7.47 percent), and those with Incoming Dark Web Traffic (7.08 percent).

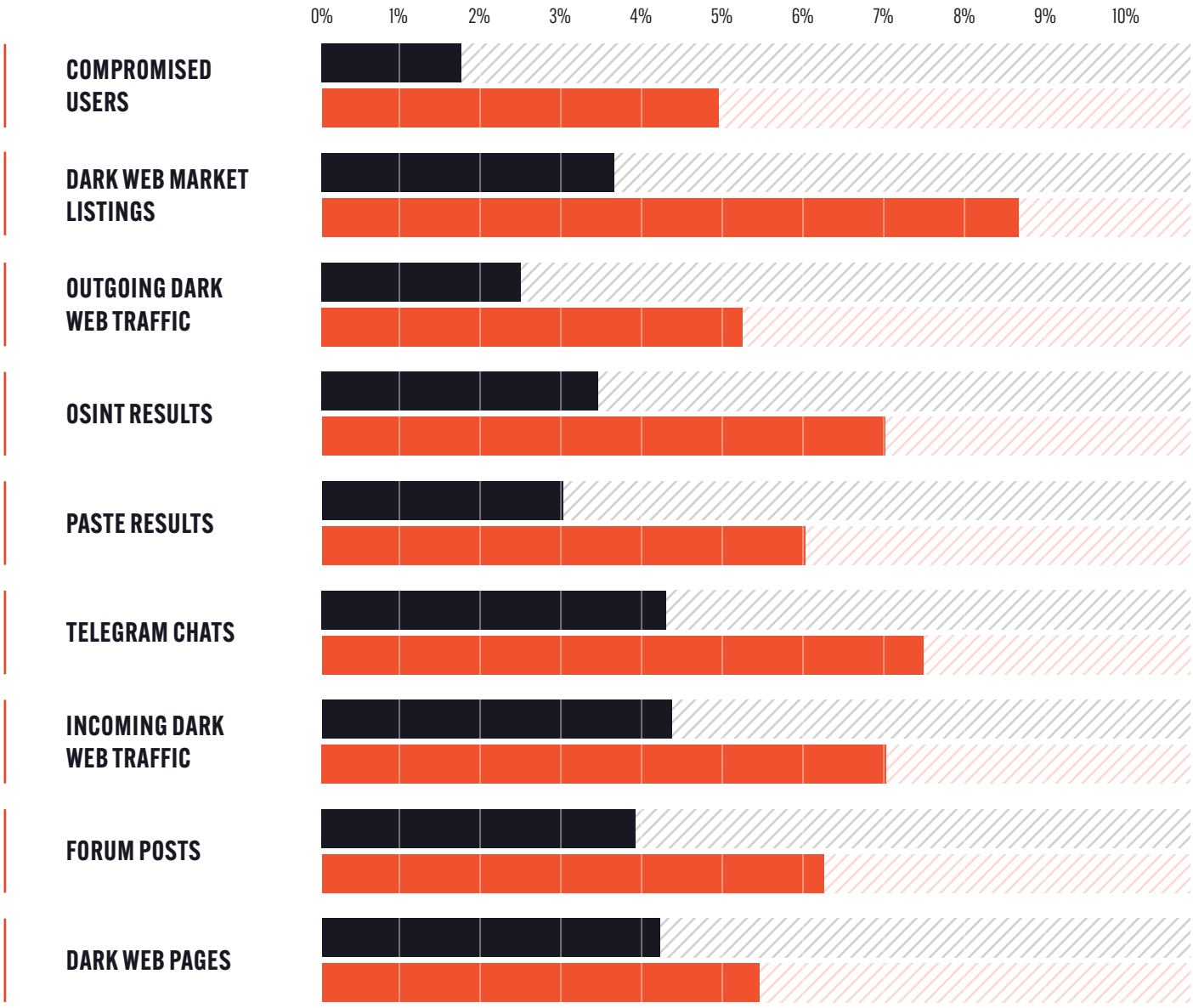


Figure 1: The impact of each dark web finding in cybersecurity risk.

Another way to look at this data is to look at the difference between the breach rate with and without findings, to determine the increased likelihood of an attack when an item is found on the dark web.

Figure 2 shows that the presence of Compromised Users has the greatest impact on the likelihood of breach, compared to when no Compromised Users were found. Four factors make the organization more than twice as likely to experience a breach than if they weren't present: Compromised Users (2.56x), Dark Web Market Listings (2.41x), Outgoing Dark Web Traffic (2.11x), and OSINT Results (2.05x).

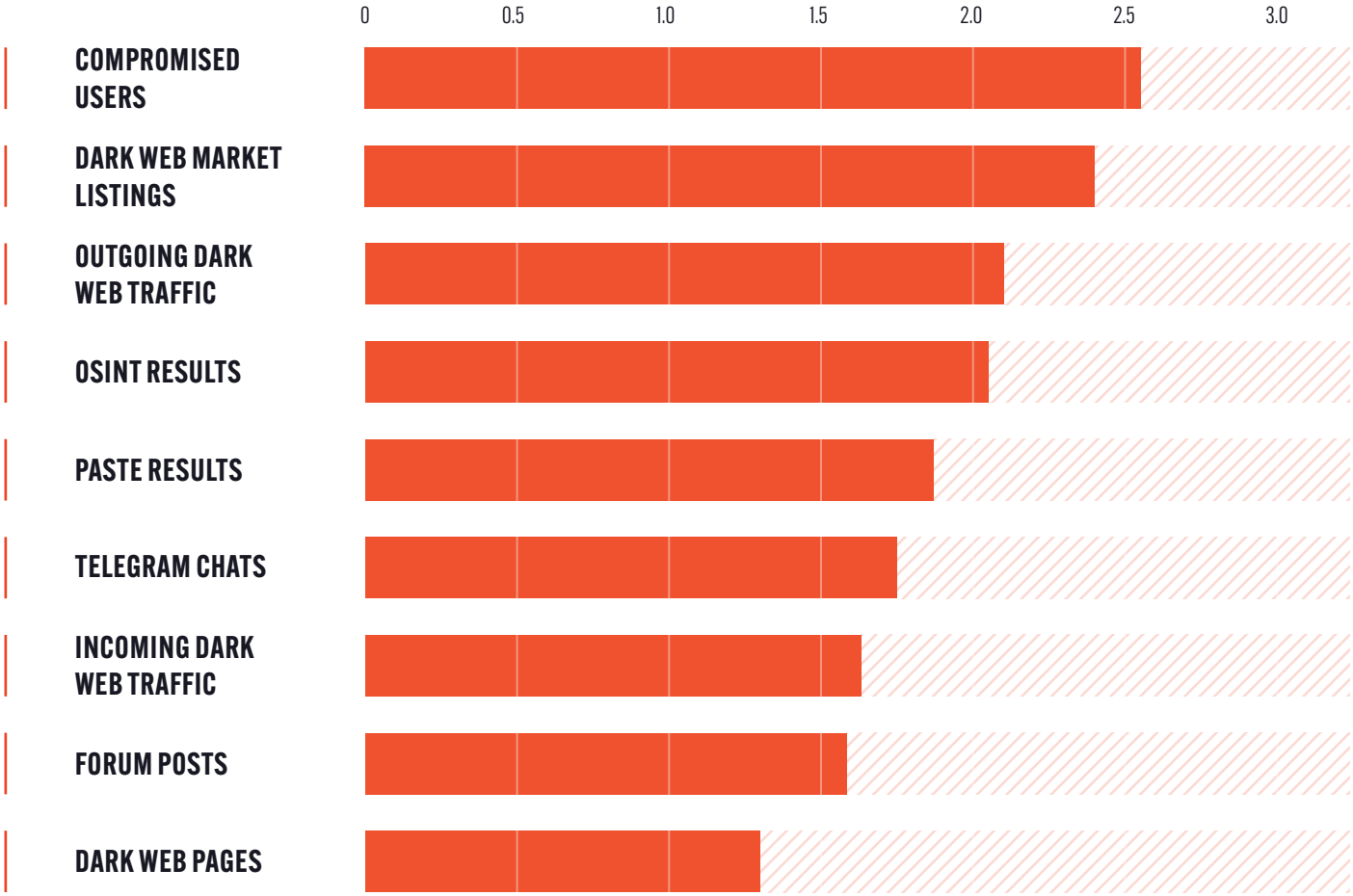


Figure 2: *The increased likelihood that organizations will have a cybersecurity incident if findings are present in these categories.*

THE INDIVIDUAL IMPACT OF DARK WEB FINDINGS

INTELLIGENCE SOURCES IDENTIFIED FOR YOUR ORGANIZATION	INCIDENT RATE WITHOUT FINDINGS	INCIDENT RATE WITH FINDINGS	CYBER INCIDENT LIKELIHOOD
COMPROMISED USERS	1.87%	4.78%	**2.56x**
DARK WEB MARKET LISTINGS	3.61%	8.69%	**2.41x**
OUTGOING DARK WEB TRAFFIC	2.47%	5.21%	**2.11x**
OSINT RESULTS	3.41%	7.00%	**2.05x**
PASTE RESULTS	3.20%	6.01%	**1.88x**
TELEGRAM CHATS	4.28%	7.47%	**1.75x**
INCOMING DARK WEB TRAFFIC	4.35%	7.08%	**1.63x**
FORUM POSTS	3.95%	6.23%	**1.58x**
DARK WEB PAGES	4.20%	5.42%	**1.29x**

Figure 3: *The results of the single-variable analysis. All dark web findings were shown to increase the likelihood of an insurance loss.*

THE COMBINED RISK OF MULTIPLE DARK WEB SOURCES

Marsh McLennan then undertook a multi-variable analysis, which is used to establish the importance of an intelligence source if an organization has results in more than one category as well as controlling for industry and revenue, which also correlates with frequency. Results in multiple categories provide a more accurate indication of cybersecurity risk but the risk model has to account for correlations between sources (see methodology in **Appendix 2**).

This analysis showed that five dark web intelligence sources are statistically correlated with risk in combination with other factors: Paste Results, OSINT Results, Dark Web Market Listings, Outgoing Dark Web Traffic, and Compromised Users.

This does not suggest that the other intelligence sources aren't predictive - as the previous section establishes, they are all independently correlated with cybersecurity risk - but ascertains which sources are most useful when multiple correlated sources are present.

Figure 4 shows that Paste Results are the most correlated with risk in a multi-variable analysis, increasing the risk of a cyberattack 1.15x. In other words, the presence of an organization's domain, email addresses, or code repositories on a paste site increases its likelihood of a cyberattack by 15 percent, in comparison with its peers.

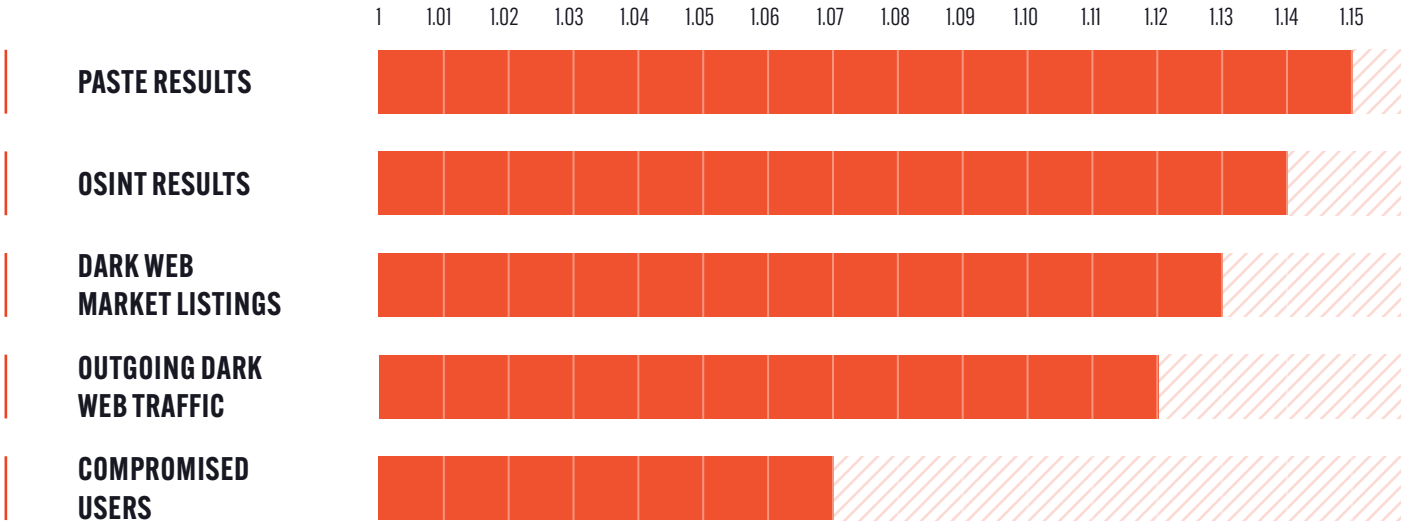
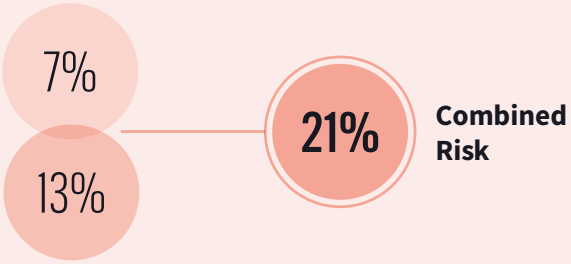


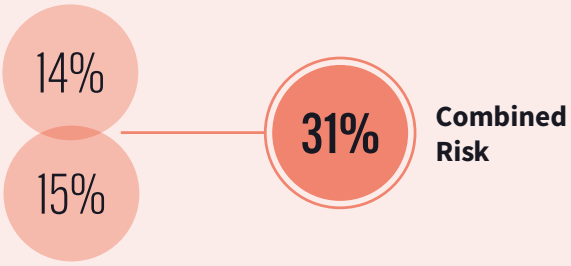
Figure 4: *The increased likelihood of a cyber incident based on a multi-variable analysis that accounts for the correlation between intelligence sources, industry, and revenue.*

The benefit of multi-variable analysis is that it allows us to accurately calculate the cybersecurity risk in the likely scenario that an organization has results for more than one intelligence source, as there is correlation between the categories (see calculation in **Appendix 2**). For example:

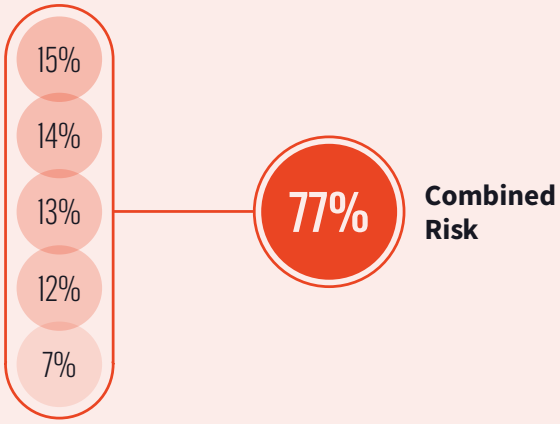
If an organization has **Compromised Users (7 percent)** and **Dark Web Market Listings (13 percent)** it is overall 21 percent more likely to suffer a cybersecurity incident relative to its peers.



If an organization has **OSINT Results (14 percent)** and **Paste Results (15 percent**) it is 31 percent more likely to suffer a cybersecurity incident relative to its peers.



If an organization has results in **all five of the intelligence sources** in Figure 4 it is **77 percent** more likely to suffer a cybersecurity incident relative to its peers.



		PASTE RESULTS	OSINT RESULTS	DARK WEB MARKET LISTINGS	OUTGOING DARK WEB TRAFFIC	COMPROMISED USERS
	INCREASED LIKELIHOOD OF A CYBER INCIDENT	**1.15**	**1.14**	**1.13**	**1.12**	**1.07**
PASTE RESULTS	**1.15**		1.31	1.30	1.29	1.23
OSINT RESULTS	**1.14**	1.31		1.29	1.28	1.22
DARK WEB MARKET LISTINGS	**1.13**	1.30	1.29		1.27	1.21
OUTGOING DARK WEB TRAFFIC	**1.12**	1.29	1.28	1.27		1.20
COMPROMISED USERS	**1.07**	1.23	1.22	1.21	1.20	

Figure 5: *The combined risk if two dark web intelligence sources are present.*

RECOMMENDATIONS

Marsh McLennan's analysis shows a statistical correlation between dark web exposure and cybersecurity risk. Searchlight Cyber recommends the following five steps organizations should take to act on these results:

1

GAIN VISIBILITY INTO YOUR DARK WEB EXPOSURE

The first step all organizations should take is to gain visibility of their exposure by collecting dark web intelligence. Cybersecurity teams can only take mitigating actions to reduce cybersecurity risk if they can identify where on the dark web they are being targeted. Understanding their points of exposure and mitigating their cybersecurity risk will help organizations minimize the financial, reputational, and legal impact of cyberattacks.

2

PURSUE COMPREHENSIVE COVERAGE OF DARK WEB INTELLIGENCE SOURCES

Each dark web source is individually a reliable indicator of cybersecurity risk but if an organization is missing sources there may be threats that they are unaware of. As the study shows, they will also have a less reliable view of their combined cybersecurity risk. Cybersecurity teams need to establish that they have coverage of all areas of the dark web - marketplaces, forums, paste sites, Telegram channels, and dark web sites. They also need to monitor for particular signals that could indicate an impending cybersecurity attack - such as dark web traffic to and from the corporate network, or compromised users that could be a path into the corporate network.

3

FOCUS ON ACTIONABLE INTELLIGENCE

While the presence of findings in any dark web category provides organizations with a high-level overview of risk, a comprehensive assessment of the threat and the implementation of defensive measures requires a far more granular level of intelligence. For example, knowing that a cybercriminal is discussing an organization on a dark web forum is not enough. Cybersecurity teams need sources that provide them the intelligence on the forum, the cybercriminal that has made the post, and what exactly the cybercriminal has posted about them in order to establish what action they need to take next.

4

USE EXPOSURE AND RISK DATA TO INFORM DEFENSIVE STRATEGIES

Visibility into where the organization is exposed, and knowledge of the risk associated with each of those threats, is most valuable when it is applied in prioritizing resources. All cybersecurity teams have stretched resources, which means that tough decisions need to be made on where budget, staff, and tooling can be applied most effectively. These decisions should be led by intelligence on the most likely paths to a cyberattack. Ideally, this intelligence should be specific to the organization, as each business has its own unique cybersecurity challenges and adversaries.

5

UNDERTAKE CONTINUOUS MONITORING

A single, point-in-time analysis of an organization's dark web exposure is insufficient for protecting an organization in the long term. As cybersecurity professionals are well aware, the threat landscape is always evolving. Cybercriminals develop new tactics, identify new points of weakness, and select new targets on a daily basis. Organizations therefore need to continuously monitor their dark web exposure for the earliest possible warning of an emerging threat. Cybersecurity risk will fluctuate over time but if organizations identify new threats quickly they will be most successful in stopping the risk of a cyberattack from becoming a reality.

APPENDIX 1: SINGLE-VARIABLE ANALYSIS

Single-variable analysis was calculated by establishing the breach rate for those who had a finding in a specific category versus those who didn't, with confidence intervals created using a beta distribution.

Figure 6 shows the single-variant analysis presented on **Page 10** with confidence intervals displayed. The fact that the confidence intervals between "Incident Rate Without Findings" and "Incident Rate With Findings" do not overlap in any category demonstrates that all dark web intelligence sources are statistically correlated to increased insurance loss frequency.

For example, findings for Compromised Users was shown to have a 4.78 percent breach rate in Marsh McLennan's analysis (indicated by the arrows) with a lower confidence interval of 4.46 percent and an upper confidence interval of 5.12 percent. What is significant is that the lower confidence interval is higher than the upper confidence interval of no findings (2.57 percent).

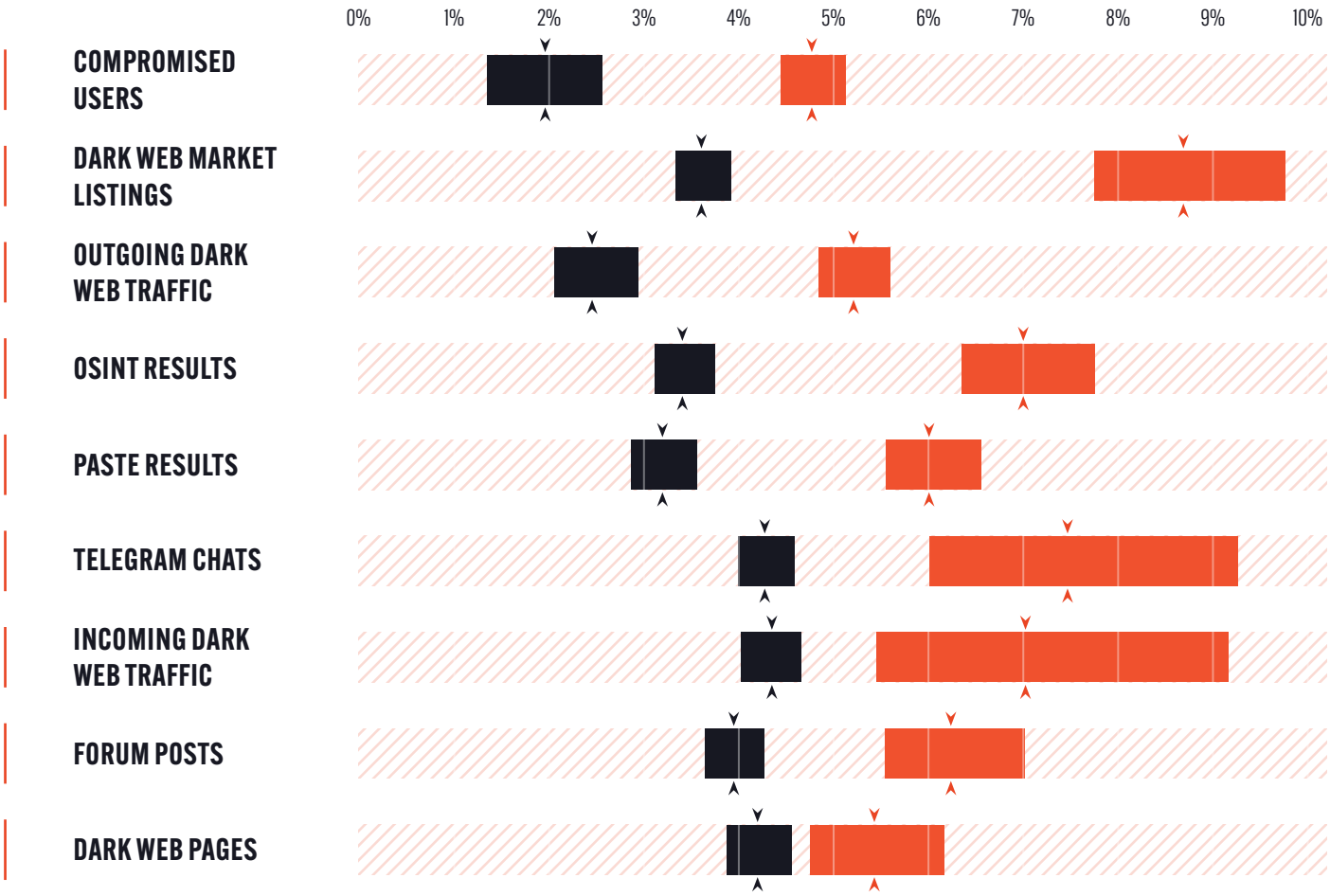


Figure 6: Single-variant analysis including confidence intervals.

APPENDIX 2: MULTI-VARIABLE ANALYSIS

Multi-variable analysis uses a classification model to measure the combined effect of multiple fields and for hidden factors like revenue and industry. This model accounts for the correlation between factors and therefore is a more reliable indicator of risk if more than one factor is present.
For example, this study showed there was high-correlation between Forum Posts and other factors, whereas Paste Results had less correlation with other factors, making it more significant as an indicator of cybersecurity risk.

For the multi-variable analysis, Marsh McLennan used a logistic regression model built with industry and revenue accounted for as controls. Confidence intervals were established by sampling the data with replacement and fitting logistic regression models 1,000 times. If the odds ratio is greater than 1 and the confidence interval does not include 1, it can be concluded that there is a statistically valid increased risk associated with a dark web finding.

Marsh McLennan's analysis placed Telegram Chats, Dark Web Pages, and Forum posts below 1.00. Even Incoming Dark Web Traffic, which McLennan found to be 1.01, was discounted because the confidence interval crossed 1.00.

The calculation for combined effect (shown on **Page 14**) is determined by multiplying the odds ratio together. For example, Compromised Users (1.072) x Dark Web Market Listings (1.129) = 1.210, or a 21 percent increase of experiencing a cybersecurity incident.

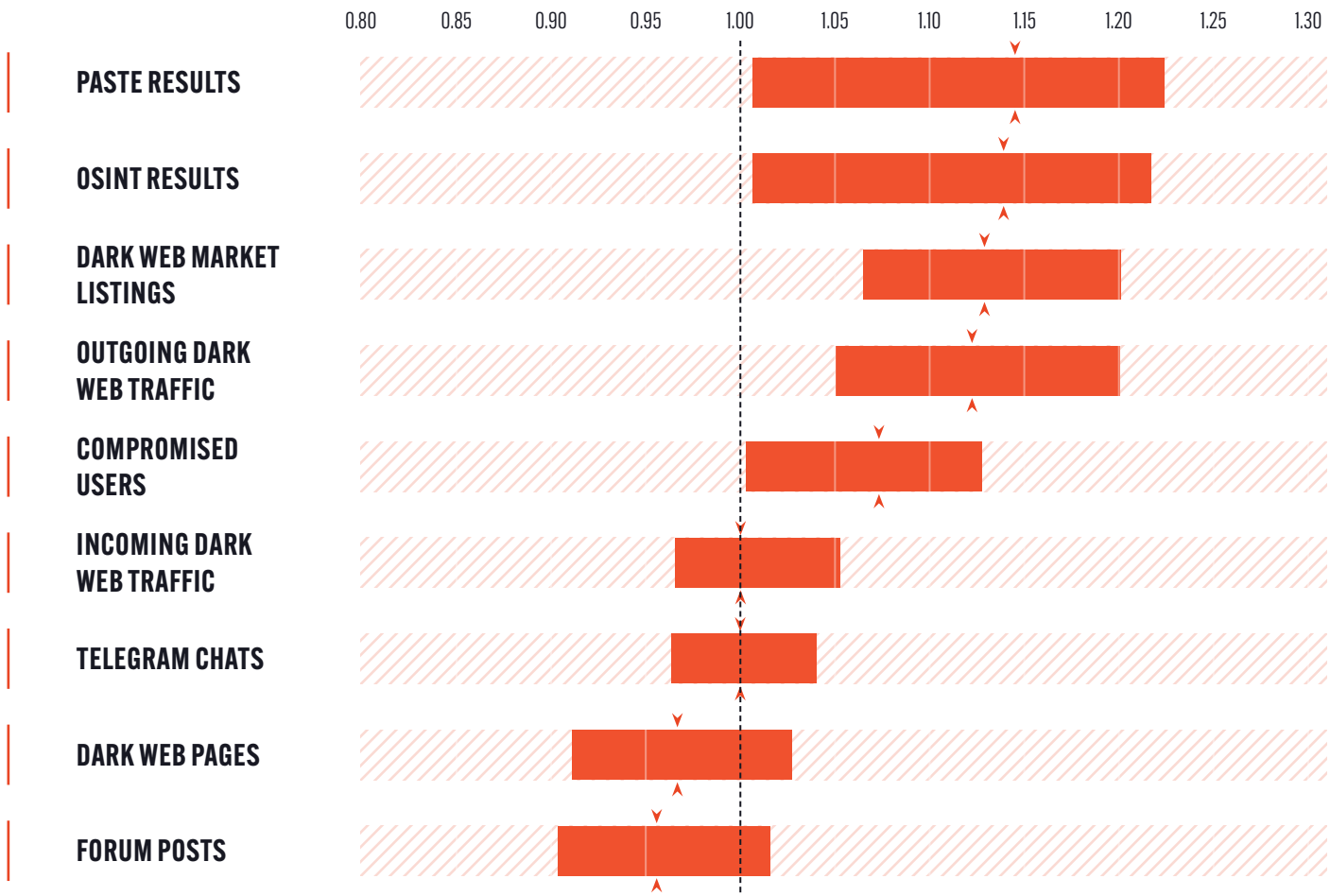


Figure 7: *Multi-variable analysis including confidence intervals.*

VISIT **WWW.SLCYBER.IO** TO FIND
OUT MORE OR BOOK A DEMO NOW.

SEARCHLIGHT CYBER

VISIT **WWW.SLCYBER.IO** TO FIND OUT MORE OR BOOK A DEMO NOW.

UK HEADQUARTERS
Suite 63, Pure Offices,
1 Port Way, Port Solent,
Portsmouth PO6 4TY
United Kingdom

US HEADQUARTERS
900 16th Street NW,
Suite 450, Washington,
DC 20006
United States