·|¦|· Recorded Future

How You Can Use the Dark Web for Threat Intelligence

Table of Contents

·|⫶|· Recorded Future

There has been much speculation (not to mention exaggeration) over recent years about the fabled dark web. We've heard how this shady underworld is the refuge of the cybercriminal elite, and even nation-state threat actors. That this is their "Wolf's Lair," where they gather to plot the breaching of businesses, the downfall of governments, and the hacking of celebrities.

As with much mainstream reporting of technology, and cyber threats in particular, there's a grain of truth here. This less accessible and more volatile corner of the internet as we know it does offer those with less honorable motives a secret marketplace for their wares.

The confusing terminology around what the dark web is or isn't shouldn't be a barrier to defenders realizing the potential benefits of information gathered from these anonymous communities, and how it can be used to produce valuable threat intelligence.

The Dark Web: A Definition

So, what's in a name? There has been a tendency to label the dark web as "any website not indexed by Google," but this definition is far too broad.

Recorded Future's Director of Advanced Collection, Andrei Barysevich, has worked as a consultant for the FBI cyber division and with international law enforcement on many cases involving Russian cybercriminals. More recently, his dark web research has uncovered breaches of government agencies and other organizations. He provides a clearer definition of the dark web:
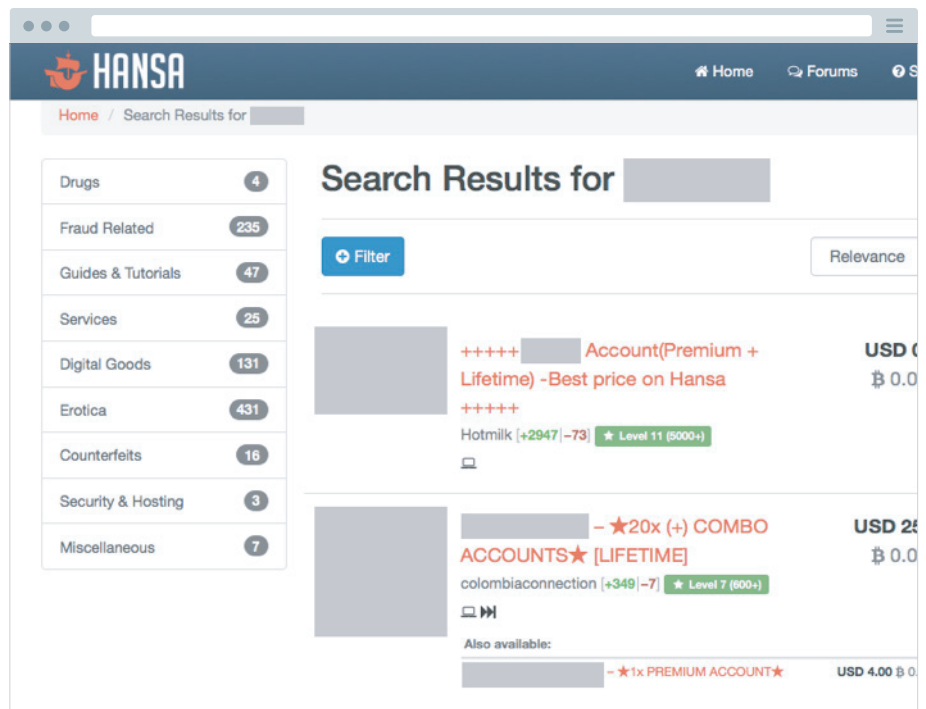
"The term 'dark web' can be confusing. I'd like to name it the criminal underground. Let's imagine a nondescript entrance to a bar in a dark alley. A place which you will not find in the yellow pages. If you know the secret knock and password, they'll let you in. Otherwise, good luck next time. The same concept actually applies to the criminal underground, or dark web communities, which you will not find via Google or any other search engine. Some of them may only be onion sites accessible through Tor, and others might only have an IP address, but no name at all."

Much has been made of the scale of the dark web, with many news outlets reporting that Google only indexes somewhere between five and 20 percent of websites. But these statistics are misleading, as the overwhelming majority of non-indexed pages are perfectly legitimate, but hidden from search engines behind paywalls or within databases.

Instead, it's best to think in terms of purely illegal activity. In this context, the dark web consists of a few hundred (maybe more) illicit communities, where members buy and sell everything from custom-made tools for cybercrime to stolen data, drugs, weapons, and more.

Another common misconception about the dark web is that only members of powerful crime syndicates use it to conduct business. In reality, the average dark web "vendor" is a small-time hacker, often called a "script kiddie," who only engages in illicit activities on a part-time basis. In fact, many members of dark web communities barely qualify as criminals at all, looking merely to purchase compromised accounts from popular services (e.g., Spotify), or discounted software.

But while the contents of dark web markets may be unavailable elsewhere, their appearance would be strikingly familiar for the majority of people. Many of the most significant dark web markets utilize a simple, bulletin-board-style interface that looks very much like Craigslist.

HANSA marketplace listings.

By contrast, the image above is taken from the popular HANSA market, which plays host to a wide variety of low-level cybercriminals, drug dealers, and fraudsters. This particular market is free to join and doesn't require any existing members to vouch for new applicants. As a result, membership is fairly high by dark web standards, likely falling in the tens of thousands or even hundreds of thousands, making the more automated eBay-style layout necessary. But not all dark web markets are as accessible as this example. The most secretive communities are highly selective about who they'll grant membership to, and often have no more than a couple hundred members.

The Deep Web Is Different

Given that there's no official classification for the deep web or the dark web, the lines between the two can get a little blurred here. Deep web usually refers to communities not indexed by search engines, that are often behind logins but probably don't require the use of Tor or onion anonymizing. While some hacker forums exist on the deep web, it also encompasses not-quite-so-nefarious information, like databases, government sites, and academic journals.

The Value of Dark Web Sources for Threat Intelligence

There's no doubt that researchers can use the dark web to obtain highly valuable threat intelligence, quite often relevant to a broad spectrum of potential targets, both organizations and individuals, otherwise not accessible through conventional monitoring.

For example, healthcare organizations can identify compromised patient records and financial institutions can analyze stolen payment information for common points of purchase, and mitigate against future fraudulent charges. In one recent case, a multinational software company prevented the sale of highly sensitive source code of yet-to-be-released enterprise software. The threat actor turned out to be an insider who was working for the company. He stole the code and was attempting to sell it on the underground for $50,000.

In another case, a government resource was penetrated by an Eastern European cybercriminal who attempted to sell the vulnerability he had identified to unfriendly state actors. Once again, the sale was identified through dark web research, law enforcement agencies were alerted, and the sale was prevented. If your organization is breached and sensitive information is stolen, there's a possibility it will show up on the dark web before you know an attack has taken place.

In these cases, breached data or intellectual property were being transacted solely via the dark web. But there are circumstances where information from open, technical, and dark web sources can be combined to deliver a much more complete view of emerging threats.

Information Seeps to the Surface

Valuable intelligence can seep to open sites from even the darkest corners of the web, the kind of information that exists in these marketplaces has a value, and that value is hard to realize unless it's seen by a wider audience. So it has to be marketed in some way.

Wikipedia broadly defines three use cases for the dark web (or darknet):

1. Out of privacy concerns or for fear by dissidents of political reprisal

2. To publish for criminal gain

3. To share media files (sometimes copyrighted material)

The argument that information needs to be made accessible outside of the dark web to realize its (monetary) value holds for both the second and third items in this list. The surface and deep web contain links to the dark web.

Recorded Future decided to explore the open, deep, and dark parts of the web to see what information is available and how it is connected. What we found was that, in terms of intelligence that can be uncovered, there really is no sharp border between them. Information tends to move into the surface web from its darker parts, meaning that the web really has many shades of darkness. The logic behind this is that brokers of illicit information on the dark web need to market their products, and so, need to post links to them on the surface web (Brian Krebs has noted the same).

Using Recorded Future's real-time threat intelligence we've identified paste sites and forums as primary nodes of communication between the surface and dark web, and can show how these are used to link to both Tor/.onion and various download sites. Our Threat Intelligence Machine™ harvests and analyzes metadata (link patterns, activity levels, and topics) from more than 800,000 sources, including technical data, open information, and the dark web. This gives us access to exceptionally valuable intelligence for threat analysis.

A Journey to the Dark Side

So, how do we see links in information between the dark and surface web in reality? Using Recorded Future to investigate, we were able to see not only how threat-related content exists on the dark web, but also how that data surfaces to other, more surprising, parts of the internet.

In this first example, we used the "TOR Uncensored Hidden Wiki index" (http://zqktlwi4fecvo6ri.onion/wiki/index.php/Main_Page) to manually locate a dubious reseller of credit cards (Premium Cards, http://slwc4j5wkn3yyo5j.onion/):
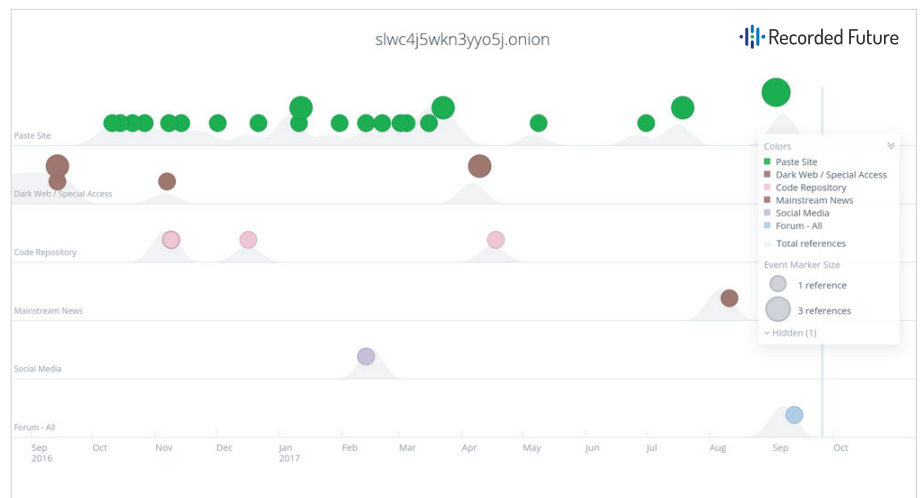
We then queried Recorded Future for the .onion link to this site and found a number of references over the last year.

You can see from this timeline that these references came from a mixture of sources. This illustrates why links to the dark web exist on both deeper, unindexed pages and up on the open web.
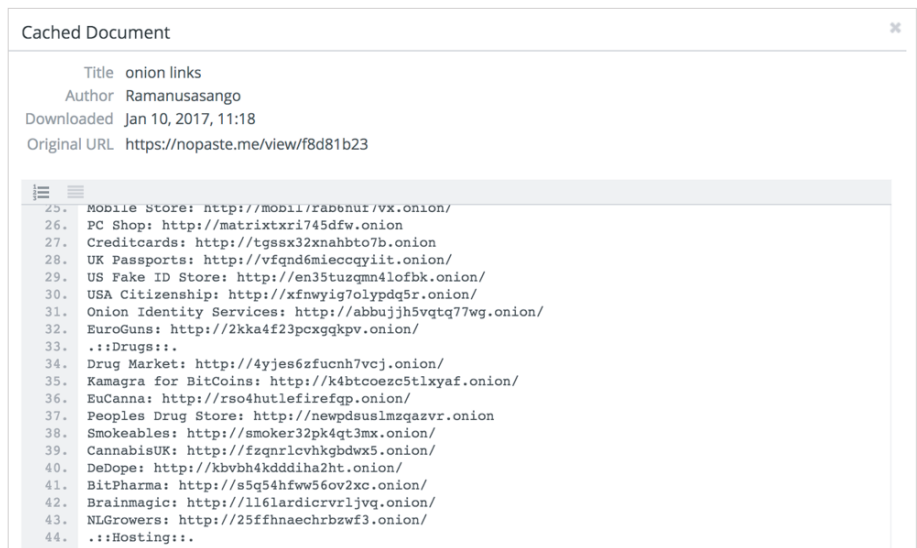
Premium Cards dark web site.

We then queried Recorded Future for the .onion link to this site and found a number of references over the last year.

A Recorded Future timeline of references and sources to the Premium Cards URL.

You can see from this timeline that these references came from a mixture of sources. This illustrates why links to the dark web exist on both deeper, unindexed pages and up on the open web.

Cached paste of links to TOR onion sites.

Here is a summary of the references and sources:

Paste Sites

Links to the Premium Cards dark web site on paste sites are primarily lists of "useful" links to sites on the Tor network. Here we can also see links to other underground marketplaces for counterfeit currency, fake passports, and illegal drugs.

Code Repository

Although GitHub is best known for code-sharing capabilities, its Gist feature lets users share any notes and snippets, too. These references are also .onion link lists.

Social Media

This tweet is from an account that regularly shares screenshots and links to sites on the dark web.

Dark Web

This reference is to a post in dark web forum by "tony.stark," a very prolific and established user. They are also questioning whether this site and another offering similar services can be legitimate.

Mainstream News

This reference is from a Reddit page called "Home of the Fraudsters." A member of the community is asking about the validity of the Premium Cards site.

Forum

A user on 4chan is asking for dark web sites where cards pre-loaded with cash can be bought and uses Premium Cards as an example.

Threat Intelligence From the Dark Web

So, the references nearer the surface we see here are very similar to those in the darker parts of the web. But further research on "tony.stark" and their posts to this dark web forum gives us an insight into the methods and targets in discussion in this particular community.

From a quick search of all "tony.stark" posts in the forum, we can see they talk a lot about spamming and phishing. On more than one occasion they advertise a unique method which guarantees a return. Banking targets seem to be a particular area of expertise.

Dark web community member offers an "original" spamming method.

"tony.stark" also recruits other forum members with expertise in coding or photoshop to help him build scam pages to link from spam emails. In return, these forum members will be vouched for and raise their esteem in the community. In this posting, "tony.stark" specifically refers to creating scam pages for Bank of Montreal and Bank of Nova Scotia.
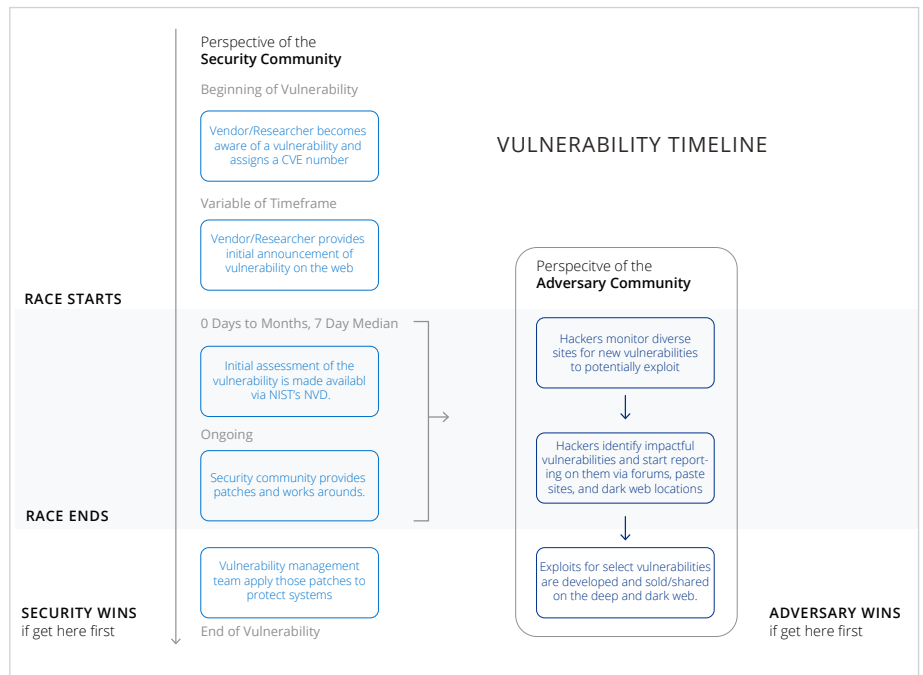
Other forum member recruited to create banking scam page.

This brief research illustrates how analysis of dark web references can uncover intelligence around attack methods and targets. It also leads us to conclude that further insight might be available from this particular dark web community, and could reveal specific conversations relevant to your organization or industry.
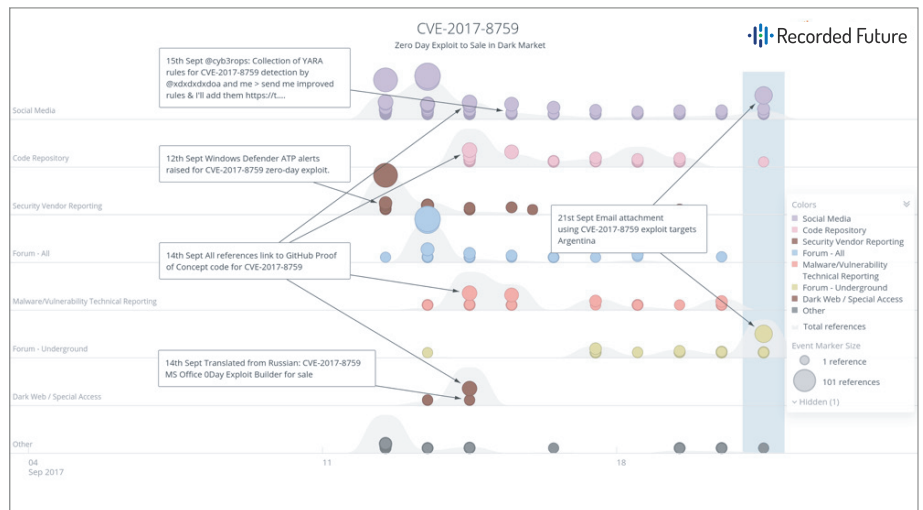
Vulnerabilities and Exploits

Recent analysis by Recorded Future revealed that 12,517 vulnerabilities (CVEs) were published on NVD (National Vulnerability Database) in 2016-2017. Almost 700 of these CVEs were first mentioned on the deep and dark web from 53 different sources. This intelligence also revealed 91 unique actors involved with some authors working with as many as 10 CVEs. Clearly, these individuals are monitoring a diverse set of sources to keep up with the most recent vulnerabilities.

The disclosure of a new vulnerability propels activity in shadier communities, and from that point, the race is between those developing and deploying the patches or the exploits. We often see reposts of researcher reports directly into deep and dark web sources. They may even be translated, for example, into Russian on Russian criminal forums. Proof-of-concept code is discussed, posted, and sold. Threat actors aren't shy about explicitly using the CVE identifier as they are doing their work. The efficiency of a naming convention is as useful for adversaries as it is for security professionals.

VULNERABILITY TIMELINE

Perspective of the Security Community

Beginning of Vulnerability

Vendor/Researcher becomes aware of a vulnerability and assigns a CVE number

Variable of Timeframe

Vendor/Researcher provides initial announcement of vulnerability on the web

RACE STARTS

0 Days to Months, 7 Day Median

Initial assessment of the vulnerability is made availabl via NIST's NVD.

Ongoing

Security community provides patches and works arounds.

RACE ENDS

Vulnerability management team apply those patches to protect systems

SECURITY WINS
if get here first

End of Vulnerability

Perspecitve of the Adversary Community

Hackers monitor diverse sites for new vulnerabilities to potentially exploit

Hackers identify impactful vulnerabilities and start reporting on them via forums, paste sites, and dark web locations

Exploits for select vulnerabilities are developed and sold/shared on the deep and dark web.

ADVERSARY WINS
if get here first

How adversaries and security teams approach newly disclosed vulnerabilities.

We can see an example of this in available intelligence around a Microsoft vulnerability disclosed in September 2017, CVE-2017-8759.
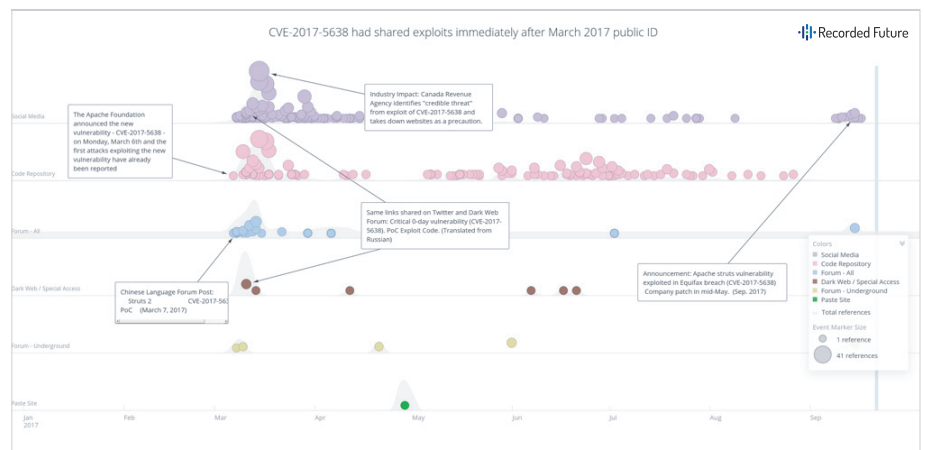

Timeline from Recorded Future illustrating disclosure of vulnerability CVE-2017-8759.

The vulnerability was a zero day already being exploited to deliver "Finspy" malware. After the disclosure, it takes only two days for proof-of-concept code to be shared across dark web communities, hacker forums, and social media. On the same day, a member of a dark web marketplace advertises this exploit for sale. In less than 10 days, news sources are already reporting this exploit attached to a phishing email, purporting to be from an Argentine government agency responsible for tax collection and administration.

This is a clear demonstration of how intelligence surrounding vulnerabilities and exploits isn't unique to dark web sources, and more importantly, how combining available information can uncover unique insights that will help you truly measure the risk. The key milestones in the evolution of this threat can all be readily revealed from disclosure of this as a zero day, right up to the sale of an exploit and its weaponization.

We can conduct this same exercise with a recent high-profile breach. CVE-2017-5638 is a vulnerability in Apache Struts, a web application framework.

Recorded Future timeline of CVE-2017-5638, used in the Equifax breach.

In this case, proof-of-concept code was being shared as soon as the vulnerability was announced. Less than two weeks after disclosure and initial proof of concepts, the Canada Revenue Agency takes the unusual step of bringing down a number of its web servers in response to what it deemed to be a "credible threat" from this vulnerability being exploited. The exploit continued to be developed, and the code shared, right up until September 8, 2017 when Equifax disclosed that data from 143 million customers had been breached, claiming that the hack happened in mid-May.

Using these kinds of indicators can equip security teams to more effectively prioritize how and when to remediate vulnerabilities. Decisions can be informed based upon real risk rather than guessing about vulnerabilities likely to be exploited, or numbers of susceptible systems.

Insider Threats

We've already seen that criminal forums and marketplaces are well-known for facilitating all types of clandestine transactions. Insider threat advertisements are also frequently used by actors promoting their illicit services on dark web sites — from retail cash-out services, to carding operations, to bank insiders facilitating theft. Many of these advertisements lie on closed-source forum sites, the kind that require extensive vetting and personas to maintain persistent access.

We can see some recent examples of conversations and advertisements related to insiders and specific organizations.

Dark web users seeking and selling insider access.

Insider threat alerting on closed forums or the dark web takes three forms. Monitoring for direct mentions of your organization or assets are the first priority, as mentions likely indicate either targeting or potential breach. Industry mentions or less specific targeting are the next avenue of monitoring, as mentions of a "UK bank" or "#x of banking accounts" attempt to cover the source of information. Finally, presence on closed-access forums allows direct interaction with threat actors, possibly retrieving samples of allegedly stolen information and materials as validation. These interactions are difficult and private, but may prove exceptionally valuable.
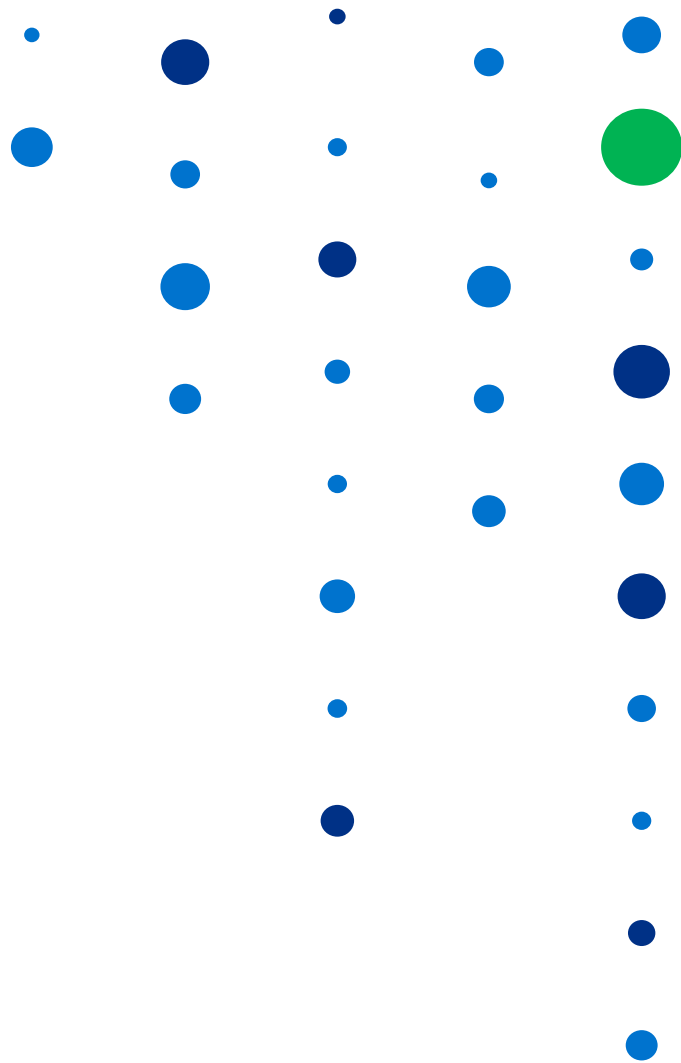
Potential Barriers to Dark Web Intelligence

We've seen the potential value of information on the dark web, but uncovering this intelligence is not without its challenges. Useful references to emerging and ongoing cyber threats can be found, but they exist among hundreds of thousands of other dark web conversations in hundreds of underground communities. This is an overwhelming amount of largely irrelevant information that risks wasting more resources rather than getting the possible benefits. Another challenge is one we've already explained: access to many of these communities is closely guarded and you may need to prove your capabilities or motivations to get into them. Finally, although the internet is global, many of the dark marketplaces exist in their own geographies and therefore, their own local languages, making understanding references to threats and targets difficult to identify and understand.

Conclusion

There is no doubt that threat intelligence gathered from the dark web is a window into the motivations, methods, and tactics of threat actors. But often, this information isn't exclusively available from dark web sources, and the most valuable insights can come by combining information across references from the surface and deep web, including technical feeds and indicators. Having the time and resources to collect, analyze, and combine intelligence from these numerous sources manually is next to impossible. There are service providers who will conduct dark web research and can deliver intelligence reports, but this does introduce the risk of a time lag, or that the information provided might not be directly related to your business, industry, or technologies.

To make the best use of dark web intelligence, you'll want to be alerted only when new and relevant information emerges, and be able to quickly determine if what's appeared requires further investigation or escalation. Using intelligence to assess risk in this way will ultimately add an extra layer of confidence in security, as well as driving more efficient decision making.

About Recorded Future

Recorded Future arms security teams with the only complete threat intelligence solution powered by patented machine learning to lower risk. Our technology automatically collects and analyzes information from an unrivaled breadth of sources and provides invaluable context in real time and packaged for human analysis or integration with security technologies.

Recorded
Future

www.recordedfuture.com

@RecordedFuture