ZEROFOX | Whitepaper

A Buyer's Guide for Digital Risk Protection

Criteria for selecting the DRP platform that best protects your digital footprint

Introduction

The wide world of digital risks

Today, public and digital risks are issues facing every organization with a digital footprint, which includes any organization that conducts business online, or promotes its brand and products on the web, or has customers reviewing or commenting on its offerings on social media platforms.

The public attack surface includes many digital risks that can be categorized as follows:

- **Campaigns of theft and fraud targeting an organization's customers,** using techniques like spoofed websites and social media accounts, phishing campaigns, and fake mobile apps. Cybercriminals use these tools to defraud the organization's customers, sell them counterfeit and pirated goods, and capture credentials used for identity theft.
- **Attacks on the reputation of an organization,** through false information on social media sites, fake reviews, and taking over social media accounts of the organization and its executives in order to post offensive content.
- **Attacks against an organization's IT infrastructure and its employees,** leading to the theft of customer data, intellectual property (IP), media content, software, employee data and credentials, and other valuable assets.

The impacts of these threats include damage to reputation and brand value, loss of customer trust, reduced revenues, regulatory fines, breach notification costs, and disruption of operations.[1]

The value of Digital Risk Protection (DRP) programs

A Digital Risk Protection (DRP) program can protect organizations from cyber, brand, and physical threats. By monitoring social media and the surface, deep and dark webs, IT security and fraud prevention teams can detect evidence of active attacks, successful data breaches and social media account takeovers, and planned campaigns by cybercriminals, hacktivists, state-sponsored hackers. They can also find indications of harmful actions by disgruntled (or merely careless) employees and customers.

An effective DRP program enables an organization to respond quickly to digital threats by taking down fraudulent websites, ads, and social media accounts, notifying review sites and online forums about false and misleading postings, and alerting online marketplaces and app stores to counterfeit merchandise and fake mobile apps. It should also provide IT groups with current threat intelligence so they can prevent data breaches by deactivating compromised user accounts, revoking stolen credentials, and strengthening security controls.

[1] For a comprehensive overview of digital risks and their impacts, see the ZeroFOX white paper "A Taxonomy of Digital Threats" at https://www.zerofox.com/resources/taxonomy-of-digital-threats/.

Criteria for selecting the right solution for your organization's DRP needs

It is extremely difficult to create an effective DRP program with manual methods. There are far too many social media platforms and websites to monitor with conventional search tools, even with a large staff of trained analysts. Critical information may be hidden in deep and dark web sites that are hard to find and access. Threat actors exchange not only in English but in languages like Russian, Chinese, Indonesian, and Arabic. Text searches are not enough, because many fraud and phishing campaigns disguise malicious content in images and video.

To address these challenges, many organizations have turned to DRP platforms and managed services. These provide the benefits of state-of-the-art tools and highly trained staffs, delivered in a scalable services model. In addition, security teams can rely on the DRP platform service provider to enhance technology and workflows as their digital footprint expands and the threat environment evolves.

But the DRP field is new, and it can be hard to distinguish between the solutions that have come on the market. This guide is intended to help by suggesting criteria you can use to select the platform that best protects your public attack surface, digital assets and brand.

We will review three topics related to the scope of DRP platforms:

1. **Breadth of social and digital platforms monitored**
2. **Protection of the complete digital footprint**
3. **Capabilities for remediation and take down**

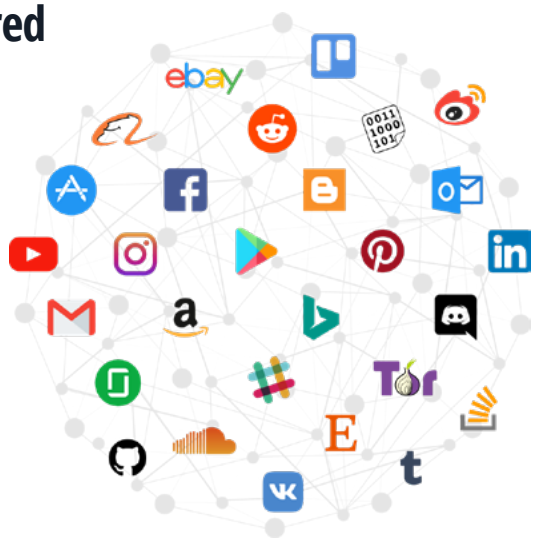
After that, we will drill down into critical capabilities in three areas:

4. **AI-based tools for advanced analysis**
5. **Global threat intelligence and research**
6. **Automation and ease of management**

DRP platforms and managed services provide the benefits of state-of-the-art tools and highly trained staffs, delivered in a scalable services model.

Criterion 1:

Breadth of Social and Digital Platforms Monitored



Information security groups have extensive investments in staff and programs for detecting Indicators of Compromise (IOCs) on their own network and systems. These are critical for stopping attacks against the organization's IT infrastructure and its employees.

But most enterprises have little experience monitoring social networks, online marketplaces, discussion forums and review websites, and deep and dark websites. They also lack tools that handle these tasks at scale. For example, when it comes to detecting threats on social media platforms and discussion forums, conventional search engines like Google and Bing:

- Are unable to discern benign from the truly bad, and hence generate massive numbers of false positives
- Provide little if any context to guide investigation into digital threats
- Have no access at all to deep and dark websites (by definition, these are not indexable or even accessible)
- Are geared for manual searches and therefore are labor intensive and do not scale well

As a result, enterprises have little visibility into the platforms that threat actors use to perpetrate frauds on customers, sell counterfeit and pirated goods, capture credentials from customers and employees, and disseminate false information about the organization, its products, and its executives.

DRP solutions are designed to give enterprises comprehensive visibility into activity on social and digital platforms. When evaluating DRP solutions, be sure to assess the breadth of their coverage. Table 1 shows the types of platforms that a DRP platform should monitor.

Table 1: Types of platforms a DRP solution should monitor

Platform Type	Examples
Global social networks	Facebook, Instagram, LinkedIn, Twitter, YouTube
Regional social networks	Sina Weibo and Tencent QQ (China), VK (Russia)
Popular websites	Glassdoor, Pinterest, Travelocity, Tumblr
Online forums and blogs	Reddit, Medium
Review sites	Amazon, Angie's List, TripAdvisor, Yelp
Paste sites	Pastebin
Collaboration / Remote work platforms	HipChat, Slack, Zoom, MS Teams
Mobile app stores	Apple App Store, Google Play, Tencent MyApp
Code share repositories	Ghostbin, GitHub, Bitbucket
Deep / dark web and covert channels	The TOR network, I2P, IRC, Telegram
eMarketplaces	Alibaba, Amazon, eBay, Mercado Libre

[1] For a comprehensive overview of digital risks and their impacts, see the ZeroFOX white paper "A Taxonomy of Digital Threats" at https://www.zerofox.com/resources/taxonomy-of-digital-threats/.

DRP solutions not only uncover evidence of ongoing attacks on social media and digital platforms, they also find threat actors discussing potential targets, recruiting accomplices, advertising wares, and exchanging information about tactics, techniques and procedures (TTPs) for conducting attacks. This type of threat intelligence is extremely useful for strengthening security controls before attacks hit, and sometimes for providing early warning of protests and physical attacks on an enterprise's facilities and employees.

DRP solutions are designed to give enterprises comprehensive visibility into activity on social and digital platforms.

Surface web, deep web, dark web: what's the difference?

The surface web: all websites that can be found by conventional search engines.

The deep web: websites and web content that cannot be crawled or indexed by search engines, including databases and sites that are protected by passwords or CAPTCHA technology. It is many times the size of the surface web.

The dark web: a subset of the deep web where communication and transactions can be carried out anonymously, typically because they use special encryption protocols and browsers such as TOR or because they require users to be invited or to pass a test. It only contains a few thousand websites, but most are difficult to find and access.

Criterion 2:

Protection of the Complete Digital Footprint

Digital risk protection is not only about where you look, but also what you look for. Enterprises need to:

- Monitor all references to their name, their brands and products, and their executives and employees, wherever they occur in social media and on the web
- Assess whether each reference is authentic, appropriate, and in compliance with regulations and corporate policies
- Detect evidence of fraud, counterfeiting, impersonation, disparagement, cyberthreats, and physical threats

DRP platforms provide tools to discover and assess references to names, brands, products, executives, and employees associated with the enterprise. Look for solutions that provide visibility into your complete digital footprint, not just parts of it. Table 2 lists the elements of a digital footprint that a DRP platform should be able to monitor.

You should also determine if a DRP solution has tools and workflows that make it easy to specify the names, brands, products, executives, logos and images, etc. that you want to monitor, and to create rules for assessing whether their use is legitimate or not.

Fake accounts and typosquatting

Cybercriminals and other threat actors often set up **fake accounts** on Facebook, Instagram, LinkedIn, Twitter, and other social media platforms that mimic the enterprise's own accounts. These accounts can be used to defraud unwary readers, capture credentials from customers, sell counterfeit and pirated products. Or they can be used to post false information or offensive content that embarrasses the organizations or its executives and employees.

Typosquatting is the practice of registering domains that are the same as legitimate enterprise domain names except for misspellings or odd characters. Typosquat domains are used to attract internet users who mistype the domain name, and as a place to send victims of phishing campaigns who may not recognize that they are not on the enterprise's website.

Table 2: Elements of digital footprint a DRP platform should monitor

Footprint Element	Monitoring and Assessment
Owned social media accounts	Taken over by cybercriminals or other threat actors? Contain inappropriate or offensive content from executives or employees? Contain information that violates regulations or corporate policies (e.g., confidential customer information, earnings reports before release)?
Non-owned social media accounts	Fake accounts used for fraud and other cybercrimes? Fake accounts used to embarrass the enterprise by disseminating false information or offensive content? Fake accounts used to sell counterfeit and pirated products?
Domains and websites	Typosquatting or homoglyphic domains registered on internet registries? Websites mimicking enterprise websites for fraud and phishing campaigns?
Enterprise, brand, and product names and logos	Names and logos used for fraud and phishing campaigns and other criminal activities? Names and logos used to sell counterfeit and pirated products?
Names/profiles of executives, employees, facilities, and events	People impersonations used for fraud and phishing campaigns? Impersonations used to embarrass the enterprise by disseminating false information or offensive content? Threats against executives, employees, and facilities?

Criterion 3:

Remediation and Takedown Capabilities

DRP is not just about detection and analysis. Those are very important, but they provide little value unless they lead to action to remediate vulnerabilities and take down web pages and social media posts used by cybercriminals, hacktivists, and state-sponsored hackers.

You should consider the degree to which a DRP platform can automate remediation and takedown activities by sending data and contextual information to SOC teams and analysts for rapid response and by facilitating automated takedown messages that can be sent immediately to social network platforms and web site hosts.

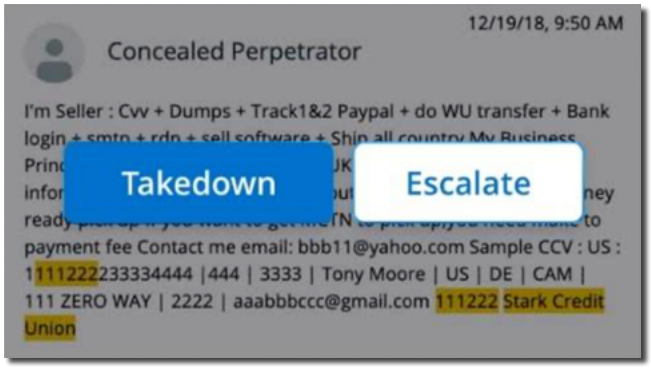


Table 3 lists some of the types of remediation and takedown actions a DRP platform should be able to initiate across the organization's public attack surface.

Table 3: Remediation and takedown actions

Footprint Element	Threats	Remediation and Takedown Options
Social media platforms	Attacks on reputation Malicious links used for fraud or phishing attacks	Request to social media platform that the post be deleted Request that the poster be blocked Request that the poster's profile be removed from the platform
Owned social media accounts	Account takeover to disseminate false information and offensive content Sensitive information inadvertently exposed by employees	Hide comments Lock account Email employees Remove unauthorized content and malicious links from all owned channels (social media accounts, owned review forums, Slack channels, email attachments)
Domains	Typosquatting and spoofed domains used for fraud, phishing, sale of counterfeit goods, etc.	Request to Internet Service Provider (ISP) or hosting provider to take down the domain Request to domain name registrar or ICANN to deregister the domain name
Surface websites	False and disparaging information Web pages used for fraud and phishing attacks Sales of counterfeit and pirated goods	Request to web sites, forums, and online markets to remove false information, prevent sales of counterfeit and pirated goods, and block users who impersonate the organization's executives and employees
Dark web forums and paste sites	Dissemination of stolen credentials, software, and intellectual property "Chatter" about planned attacks on the organization	Work with the paste sites to remove stolen credentials, software, intellectual property, and discussions of illegal activities Notify asset owner so they can harden infrastructure, revoke credentials, change passwords, etc.

Successful remediation and take down depends not just on technology. Each individual social media platform, forum, market, and web hosting site has its own contact point, takedown request process, terms of service, and requirements for evidentiary information. You need to make sure your DRP service provider has the knowledge and experience to handle takedowns for the widest possible range of platforms.

A DRP platform should also provide excellent reporting and performance metrics for takedowns, so you can track takedown requests, status, successes and failures, duration, and history.

Example of rules regarding reporting violations to social media platforms

Reporting violations of social media platforms Rules and Terms of Service

Social media platforms have separate report forms for each type of violation of their rules and terms of service:

- **Unauthorized trademark use**
- **Unauthorized use of copyrighted materials**
- **Sale or promotion of counterfeit goods**
- **Privacy policy towards children**
- **Child sexual exploitation: Learn more about our child sexual exploitation policy**
- **Pornography**
- **Impersonation of an individual or brand**
- **Private information posted on social media platform**
- **Abusive behavior and violent threats**
- **Spam and system abuse**
- **Violation of social media platform Ads policy**

Information required to report a trademark violation to social media platforms

- **Trademark holder's name**
- **Trademark holder's address**
- **Trademark holder's country**
- **Trademark holder's website**
- **Trademark holder's social media platform username (optional)**
- **Trademarked word, symbol**
- **Registration number**
- **Registration office**
- **Direct link to trademark record or trademark search page (optional)**
- **Social media platform account being reported**
- **Description of the confusion with your trademark**

Criterion 4:

AI-based Tools for Advanced Analysis

Unfortunately, digital threats are now so numerous, so diverse, and so cleverly concealed, that mere human beings can no longer hope to find and identify them all, even when aided by conventional search engines.

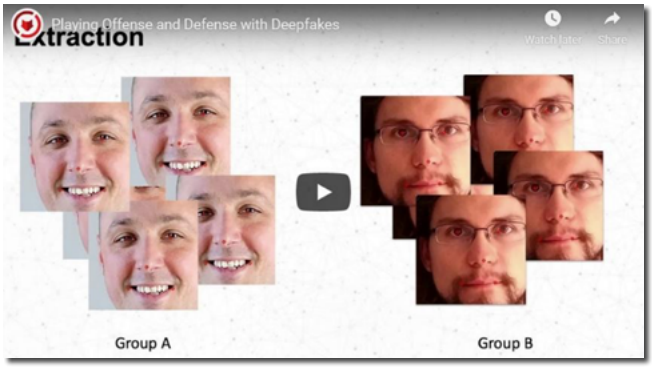
DRP platforms are incorporating advanced artificial intelligence (AI) technologies to overcome obfuscation and process huge volumes of data. You should look for DRP solutions and vendors that utilize as many of these advanced technologies as possible to provide agility, scale and offload burdensome analysis tasks from human experts.

Table 4 summarizes some of the types of AI that can be used to uncover evidence of digital risks.

Table 4: Artificial intelligence technologies and their application to DRP

AI Technology	Key Capabilities
Natural language processing (NLP)	Analyzes text in multiple languages and derives meaning and intent.
Malicious link analysis	Examines content in social media posts, collaboration tools, email communications, and other channels and detects suspicious links related to malware, phishing, spam, and other threats.
Optical character recognition (OCR)	Detects characters embedded in images used to mislead viewers, convey false information, and perpetrate fraud.
Object detection/ computer vision (CV)	Discovers patterns in large volumes of text and images that can be used to detect and classify digital threats.
Machine learning (ML)	Compares facial images to detect unauthorized use of pictures of executives and employees and impersonations.
Deepfake detection	Identifies digitally altered images and videos.
Analysis at scale	Ability to evaluate billions of content pieces annually and process millions of data points daily to enrich machine learning and other AI techniques.

Web presentation: Playing Offense and Defense with Deepfakes



Deepfake detection is one of the most exciting new applications of AI for DRP. Click on the image for an overview of the topic, and an introduction to Deepstar, ZeroFOX's open source contribution to Deepfake detection.

Criterion 5:

Threat Intelligence and Research

Threat intelligence is information about current threats and ongoing attacks, as well as about the tactics, techniques, and procedures (TTPs) of threat actors. Good threat intelligence helps InfoSec and fraud teams identify digital threats faster and respond quickly with appropriate measures.

Some DRP platforms do not offer any threat intelligence, or offer only raw data or partially digested information. You should compare DRP solutions that provide validated, prioritized intelligence, curated by experienced analysts who can help interpret the information and deliver informed analysis. Table 5 lists some of the attributes of threat intelligence that make it effective for DRP activities and deliver it in ways easily consumed by your organization.

Digital Risk Protection Makes Cyber Intelligence Accessible to the Midmarket

"Organizations in the midmarket struggle with threat intelligence even more than their enterprise peers because they have smaller budgets, fewer security personnel, and significantly less time to dedicate to the cause…Digital risk protection (DRP) is a critical step to satiating that need as it provides extremely actionable and relevant intelligence…DRP services are a great place to start because:

- DRP offerings embrace services to reduce need for a dedicated threat intel capability.
- Investments in DRP services are modest.
- DRP provides tactical intelligence, which makes it easier to demonstrate ROI."

Excerpted from the Forrester Research report The Digital Risk Protection Market In 2019: The First Cyber Intelligence Capability You Should Invest In, by *Josh Zelonis and Trevor Lyness, April 19, 2019, with permission from Forrester Research.*

Table 5: Attributes of effective threat intelligence for DRP

Attribute	Description
Contextual	Threat intelligence relevant to your use cases and threat environment.
Prioritized	Information prioritized by criticality of the threat.
Actionable	Alerts with recommended actions and 'in-alert' response options, packaged for semi- or fully-automated response.
Curated, strategic	Finished intelligence about active and emerging threats, with analysis and recommendations by expert research staff.
Option for custom threat analysis	Option to engage a dedicated analyst for custom investigations and special reports.
Flexible communication	Options to communicate threat intelligence to analysts and security operations center (SOC) operators via online portals, mobile apps, and email updates.
Integration with security tools	Ability to export threat intelligence and indicators of compromise into SIEMs, threat intelligence platforms (TIPs), security orchestration, automation and response (SOAR) tools, and other security solutions.

Criterion 6:

Automation and Ease of Management

DRP platforms vary widely in how well they automate workflows and how easy they are to use and manage. Evaluate DRP solutions to see if they automate and streamline essential processes and improve the experience of your analysts and digital security teams.

Table 6 lists some of the features of a DRP platform that facilitate automation and easy use and management.

Table 6: DRP solution features for ease of use and process automation

Feature	Value
Direct connection via APIs to social networks, web hosting companies, internet registrars, and other platforms	Enables fast, accurate collection of content, data, and context without manual effort. Enables fast submission of remediation and takedown requests.
Simple, flexible creation of rules and policies for data collection and alerting	Rules covering common use case situations should be readily available out-of-the-box pre-integrated. Rule customization should be straightforward to handle unique cases and to fine tune existing rules for accuracy.
Best-practice user interface	Increase analyst productivity with simplified navigation, filtering of menus and displays, auto-form completion and well-designed screens for tasks such as entity configuration, policy and rule creation, alert viewing, research, and the creation of remediation and takedown requests.
At-a-glance summary dashboard	Enable analysts, managers, and executives to see summary digital risk health data, status, and trends and to drill into summary data for details such as top alerts, top attacked entities, data sources, and resource highlights.
Flexible communication	Allow for real-time, anywhere access and interaction with the system for situational awareness and response via a mobile app, an online portal and email.
Flexible, role-based reporting	Provide out-of-the-box reports tailored for roles such as SOC analyst, marketing or brand manager, location security manager, and executive, with ability to produce customized reports on topics such as alerts and takedowns.

How to Select the Right Platform for Social Media and Digital Risk Protection

You should now have a solid understanding of the value of digital risk protection programs and how effective DRP platforms cast a very wide net to detect and remediate risks to your organization's reputation, revenue, brands, employees, and customers.

We hope this guide will help you in your selection to ask the right questions about the DRP solutions and DRP vendors you evaluate. To recap some of the main points:

- **Assess the breadth of social and digital platforms monitored by the DRP solution,** including global and regional social networks, popular websites, online forums, blogs, and review sites, mobile app stores, and code paste sites.
- **Determine whether the DRP solution will protect your complete digital footprint,** including owned and non-owned social media accounts, domains and websites, brand and product names and logos, and your executives and employees.
- **Examine the DRP solution's remediation and takedown capabilities,** in particular its ability to efficiently initiate and quickly complete takedowns and appropriate resolution steps with social media platforms, owned social media accounts (in case of account takeover), domains, and surface, deep and dark web forums and websites.
- **Review the DRP solution's use of artificial intelligence-based technologies** such as natural language processing, machine learning, malicious link analysis, object detection, image comparison, and deepfake detection.
- **Verify that the DRP solution or provider is able to produce and use contextual, prioritized, actionable threat intelligence** and provides options for curated strategic intelligence and custom investigations.
- **Evaluate the DRP solution for automation and ease of management features** such as API connections to social media and web platforms, simple rule creation, an excellent user interface, and an at-a-glance dashboard with drill-down capabilities.

ZEROFOX

For more information please visit www.zerofox.com

About ZeroFOX

ZeroFOX, the global category leader in public attack surface protection, safeguards modern organizations from dynamic security risks across social, mobile, surface, deep and dark web, email and collaboration platforms. Using diverse data sources and artificial intelligence-based analysis, the ZeroFOX Platform identifies and remediates targeted phishing attacks, credential compromise, data exfiltration, brand hijacking, executive and location threats and more.

The patented ZeroFOX SaaS technology processes and protects millions of posts, messages and accounts daily across the social and digital landscape, spanning LinkedIn, Facebook, Slack, Twitter, Instagram, Pastebin, YouTube, mobile app stores, the deep & dark web, domains, email and more. Led by a team of information security and high-growth company veterans, ZeroFOX has raised funding from NEA, Highland Capital, Redline Capital, Intel Capital, Hercules Capital and others, and has collected top industry awards such as Red Herring Top 100 North America, the SINET16 Champion, Dark Reading's Top Security Startups to Watch, Tech Council of Maryland's Technology Company of the Year and the Security Tech Trailblazer of the Year.