# Data Security: The Key Lies in Knowing What You Don’t Know

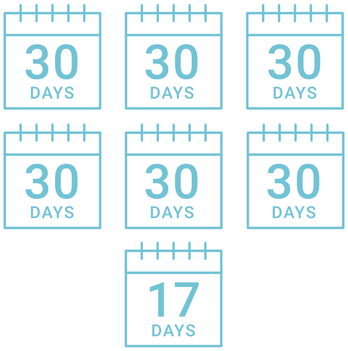
## Defending Enterprise Data Through Continuous, Real-Time Visibility and Analysis

WHITEPAPER

**Ultimately, every cybersecurity strategy and tactic is implemented to protect data.** Ultimately, every cybersecurity strategy and tactic is implemented to protect data. Today, organizations have secured endpoints, networks, applications, and online services—not 100%—but to the best of their abilities. However, data is still lost or, more commonly stolen, every day. That’s because most organizations do not have enough insight into their data. There is no data “electrocardiogram” to display normal and abnormal usage patterns. With thousands of users, petabytes of data, millions of files, and most of it in constant motion, companies can only protect what they think they know.

This paper will examine how traditional approaches to securing data—such as DLP and encryption—are falling short and why organizations are increasingly challenged by significant industry trends to defend their data. The paper also describes a new way of identifying and displaying the organization’s data assets and updating the “normal” baseline—all in real time. CrowsNest™ from Flying Cloud Technology provides the first real-time data defense solution that delivers immediate visibility into organizations’ data, its movement, usage, and changes so that it can be protected.

# The Shift Towards Detection and Response



**197 Days**
The average time to *identify* a data breach

30
DAYS
30
DAYS
9
DAYS

**69 Days**
The average time to *contain* a data breach

Data represents a huge percentage of an enterprise's value. Yet, in spite of organizations' large investments in security solutions, data is still being compromised. Data loss continues to happen all too frequently. Bad actors still steal Intellectual Property (IP), Personally Identifiable Information (PII), and other sensitive information. In spite of powerful security solutions for network perimeter and host systems, it's becoming clear that prevention measures can only reduce the onslaught of successful attacks.

Cyber attackers will get in. In many cases, they already are inside, and they're after valuable data. That can mean pre-patent product designs, proprietary manufacturing techniques, software code—whatever an organization considers its "crown jewels." To get there, attackers use every bit of data that they can find to unlock system access, masquerade as legitimate users, and exfiltrate information. That's why the focus must shift to securing data itself.

Organizations must increasingly adapt their security setup to focus on detection, response, and remediation[1]. According to the Ponemon Institute, the average time to identify a data breach is 197 days, and the average time to contain a breach once it has been identified is 69 days[2] .

Traditionally, data loss prevention (DLP) software has been seen as the primary method to prevent sensitive data loss. Yet, even the best DLP solutions can't keep pace with all of the data at risk. Organizations create DLP policies based on the data that they know they have and typical usage patterns. But new data is always being created. Existing data moves to different locations and changes as it flows nonstop through the enterprise. Where did it go? Who touched it? How? More important, why? How is usage of a specific file connected to a seemingly unrelated event elsewhere in the network?

That's the fundamental problem. Those answers are unknown, and organizations can't effectively defend the unknown. It's precisely the problem that Flying Cloud Technology solves.

[1] 5 Trends in Cybersecurity for 2017 and 2018, Smarter with Gartner, June 14, 2017

[2] 2018 Cost of a Data Breach Study, Ponemon Institute, sourced by IBM, July 11, 2018

# Guessing, Wondering, and Anticipating Isn't Enough

Every business needs a data-centric security strategy to prioritize datasets for mitigation of growing business risks caused by data protection and privacy laws, hacking, fraud, and ransomware.

Hype Cycle for Data Security, Gartner July 24, 2018

Perimeter and prevention-based security measures are primarily designed to protect users and systems. The assumption was that if cyberthreats can be kept out, data would remain safe. As a result, endpoint protection, antivirus and anti-malware solutions, email threat defenses, firewalls, fraud defense, and web filtering products are deployed as siloed capabilities. They defend against known external threats targeting specific types of network traffic and communications. Enterprises deploy encryption and DLP to prevent known sensitive data from leaking out of the enterprise.

Each of these solutions is important, and each protects within its designed capabilities. All require enterprise security teams to define policies for maximizing their value. Security teams make their best guesses about potential threats when designing and deploying policy, hoping to cover systems and users against all possible threat vectors. Yet, none of these solutions adequately protect data itself, because no one really knows what normal data behavior and usage looks like across the organization.

## Protecting What Isn't Seen

Cloud-based applications and storage create even bigger challenges for security teams. How do they write policies to secure data that they don't even know exists? Developers build applications and store data on AWS, Microsoft Azure, and Google Cloud. Marketing teams share content with agencies and external partners via Box and Dropbox. Public, private, or hybrid—the average enterprise uses 1,427 distinct cloud services[3] and 18.1% of files uploaded to cloud-based file sharing and collaboration services contain sensitive data. A proliferation of cloud services changes the game when it comes to describing "normal" data usage.

## The Encryption Dilemma

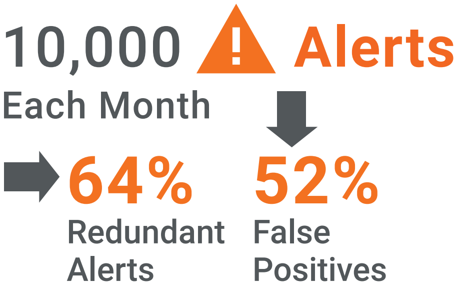
Encrypting sensitive data increases security, but encryption is complicated, costly, and not fool-proof. An encryption approach assumes that only external threats are interested in sensitive data and that employees and insiders are "safe." In today's world however, most enterprises must send, receive, and safeguard sensitive data from customers and third-party partners. Is a third-party partner an "insider?" If they must decrypt the data to use it for business purposes, who is to say that it is still secure in an environment over which the sender has no control?

Worse, what if a vendor of trusted solutions has malicious ecosystem partners? In one case, a healthcare organization used an encryption vendor for its email. A number of the encryption vendor's partners took advantage of their knowledge to deliver malware to the vendor's customers. They sent encrypted email messages to the healthcare organization. When messages arrived at the organization's email gateway, they could not be filtered because they were encrypted and therefore, assumed to be secure. When messages hit the encryption vendor's gateway next, they were decrypted and delivered—carrying dangerous malware and infecting systems across the healthcare organization.

If an organization does want to encrypt data, determining where encryption starts and stops adds exponential complexity to a dynamic environment with thousands of applications, systems, and users. Terabytes or petabytes of stored data must be scoured to find and classify sensitive information for encryption. Encrypting data in motion and in use also requires accompanying decryption capabilities—without creating such a nuisance for users that they eventually ignore policy altogether. Encrypting files that reside in the cloud is even trickier, as many clouds and SaaS providers or applications require access to unencrypted files or data at rest.

[3] 12 Must-Know Statistics on Cloud Usage in the Enterprise, Skyhigh Networks/McAfee, 2018

**Guessing, Wondering, and Anticipating Isn't Enough** (cont.)



10,000 ⚠ Alerts

Each Month

64% Redundant Alerts

52% False Positives

## The Enemy Inside

Insider threats have become pervasive. Employees have access to system credentials and often can operate across the network as legitimate users without triggering alarms. Credential thefts by company insiders have tripled in the past two years[4] with an average cost of $648,845 per incident. Out of 29 cases of IP theft executed by foreign beneficiaries, the CERT Insider Threat Center found that all of them involved malicious insiders who "misused a company's systems, data, or network to steal IP." Insider threats are especially damaging because usually by the time they are discovered, volumes of valuable data and IP are already in the wrong hands.

## Alarm Fatigue

Even effective data encryption, DLP, and other solutions still overwhelm security teams with alerts. A survey by FireEye polled C-level security executives at large enterprises worldwide and found that 37% of respondents receive more than 10,000 alerts each month. Of those alerts, 52% were false positives and 64% were redundant alerts[5]. The Cisco 2017 Security Capabilities Benchmark Study found that, due to various constraints, organizations can only investigate 56% of the security alerts they receive on a given day. Half of the investigated alerts (28% ) are deemed legitimate and fewer than half of those are actually remediated. The threat landscape is so enormous and changes so quickly that it has become impossible to pay attention to—let alone analyze—thousands of security alerts to know when data is being compromised, misused, or exfiltrated.

## Is Anything "Normal" Anymore?

Well, yes and no. Mergers and acquisitions (M&As), complex third-party ecosystems, and continuous IT adaptation to business drivers have become normal parts of enterprise life. However, protecting data assets and infrastructure in light of the new normal has become incredibly more complex. M&A activity often opens security gaps as infrastructures with differing security measures are merged. It takes time to extend comprehensive policies to completely cover the new organization. Large extended enterprises must share data with third parties over whom they have no control. Organizations must protect legacy business-critical systems, such as mainframes or wire transfer systems that were not originally designed for today's threat environment but are still at risk.

No matter how many staff, security experts, and resources are hand in an organization, it's all but impossible to monitor and assess live data in this environment. To defend its data, an enterprise must know where it resides, whether it has been tampered with, or if it's in the possession of an unauthorized user—in real time without false positive alerts.

[4] 2018 Cost of Insider Threats: Global, Ponemon, April, 2018.

[5] False positives still cause threat alert fatigue, CSO, May 3, 2017

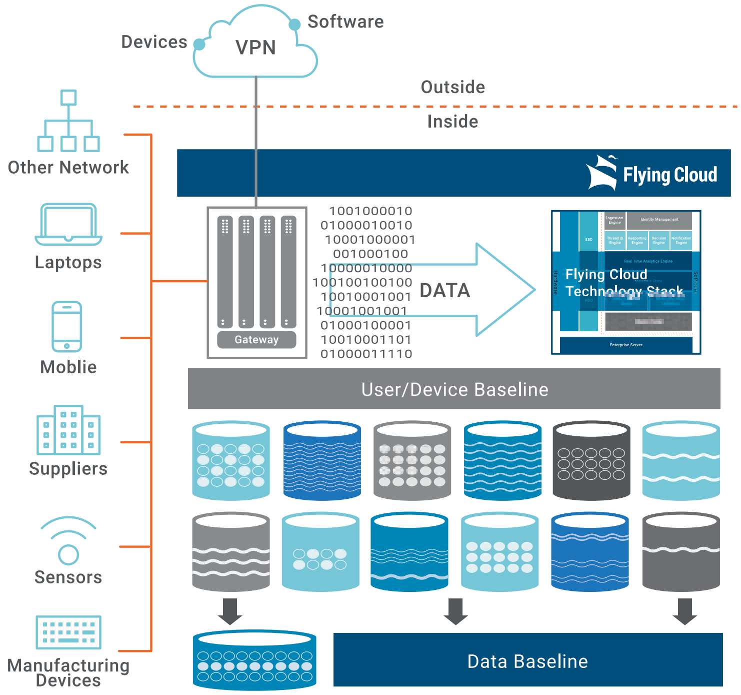
# Defending Data by Knowing the Previously Unknown

The first step in data defense is knowing your data—where it comes from, its structure, content, purpose, level of sensitivity, and relationship to other data, users, and movement. Data defense also requires an understanding of who uses data, how they use it, and what they are allowed to access. This baseline knowledge is critical. Traditional security measures rely on humans to assess threats and design behavior-based countermeasures.

Instead, CrowsNest from Flying Cloud Technology enables organizations to make the unknown, known. The first real-time data defense solution, CrowsNest delivers immediate visibility into the organization's data movement, usage, and changes. Once the organization's data and patterns are known, organizations can quantify risk and base security policies strictly on the data itself—not unpredictable behaviors.

CrowsNest analyzes incoming data, data in motion across the network, and data leaving the environment to identify and help prevent data tampering, loss, and exfiltration. It creates a rolling baseline of normal data patterns with contextual analysis, so that anomalous usage, unprivileged access, and threat actors become immediately visible. When anomalies appear, CrowsNest delivers real-time data forensics and analytics. Security defenders receive a data "chain of custody" that identifies exactly who, where, when, and how content was accessed, modified, or distributed.

## Rolling Baseline Technology™



## Use Case: Define Normal

Prevention solutions typically rely on content inspection and log data to uncover potential threats. Clever attackers can easily manipulate content to remove or hide sensitive data. Scanning or monitoring solutions often are limited to scanning stored data or data residing on certain servers. A file residing on a server might be monitored, but copies of that file are circulating around the network, cloud-based systems, and users' computers where the risk is much higher. With many monitoring solutions, data outside of storage or designated systems can't be "seen" and for defense purposes, doesn't exist. Still other solutions track files already known to contain sensitive data, but miss all other data activity.

CrowsNest monitors and analyzes "live" data—data as it moves across the network, is used, and sent. Real-time monitoring, patented machine learning, and automation quickly establish a baseline of normal data patterns and then continuously update a rolling baseline of normal activity.

With the knowledge of normal, enterprises can fast-track new DLP and other security deployments with CrowsNest analytics and reporting. Existing solutions and policies can be tuned and updated to remediate gaps based on new visibility into previously unknown data patterns.

## Use Case: Detect IP Threats

Traditional threat prevention solutions often miss odd data usage patterns or changes to files that could signal a threat. Most detect sensitive data based on previously defined keywords or textual content, using basic pattern matching techniques, which consistently result in high false positive rates.

CrowsNest redefines detection of sensitive data by analyzing its content and context against the rolling baseline. It knows normal data patterns on the network, enabling it to track where critical data is flowing and immediately detect any deviation in pattern or use. CrowsNest also validates data content and structure against other data and derivative works—such as copies of files. Data modifications, file structure changes, and unusual usage trigger further analysis and reporting.

## Use Case: Find and Contain Insider Threats

Enterprises face significant risk to data through their employees, whether intentional or not. Mobile devices containing valuable data are lost or stolen. Employees decide to jump to a competing firm. Some launch side businesses based on their employer's IP. Teams move between projects and continue to have access to data that they no longer need.

Frequently these activities go unnoticed until suspicion is triggered through other tools that notice large unusual file downloads or email attachments being sent to non-business locations. CrowsNest machine learning delivers visibility into normal data usage for individuals, roles, groups, business divisions, or any combination. When data not normally used by a group is suddenly harvested, CrowsNest detects and alerts. Reports provide comparative baselines and data that enable security teams to further investigate and prevent sensitive data modification or loss.

## Use Case: Defend Data in Shared Environments

Enterprises often share data with third parties, such as suppliers, research partners, and outside agencies. How third parties are allowed to use data is governed by contractual obligation or regulation. With multiple channels available for accessing enterprise data—direct access to internal systems, cloud access, shared drives—it becomes tremendously challenging to detect irregularities in data access or use. CrowsNest enables continuous, real-time monitoring and analysis of data shared with third parties. It also ensures effective data segregation, giving people access to the data they need in an appropriate manner.

CrowsNest provides a data "chain of custody," identifying who, when, where, and how data was used. This chain of custody for data is invaluable for legal discovery, criminal investigation, compliance, and verifying third-party agreements.

## Use Case: Accelerate Breach Response and Remediation

Security teams are already struggling to prioritize and respond to a relentless stream of threats, incidents, and alerts. Identifying a breach can still be a slow process, allowing advanced attackers to remain in a network for months without being found. Even after a breach is discovered, teams frequently lack adequate forensics to know what was exposed and how to effectively respond.

CrowsNest rolling baseline and data forensics features give analysts continuous visibility into and analysis of data movement. Anomalies are spotted as they occur, with accompanying detail so that security teams can quickly detect, effectively respond, and appropriately remediate. CrowsNest can send alerts to an organization's SIEM, ticketing, or orchestration systems through CrowsNest APIs, enabling teams to automate threat response and remediation to the degree that they want.

# Defend Data and Preserve Company Value

An organization's data represents its most valuable asset, aside from its employees. It's so important that Boards, investors, and the U.S. Securities and Exchange Commission recognize that a company's data materially affects its market and opportunity value. Now, organizations can know exactly what their data assets encompass and how they are used. CrowsNest from Flying Cloud delivers deep visibility and analysis into live data across the organization for its protection and the preservation of enterprise value.

**Learn more about how CrowsNest from Flying Cloud Technology**
can protect your organization's data. Visit **flyingcloudtech.com**.

**About Flying Cloud Technology**
Flying Cloud Technology provides innovative data defense solutions that enable enterprises to protect IP, PII, and other essential or sensitive data. The company's CrowsNest platform uses advanced machine learning and big-data analysis, enabling customers to identify cyberthreats, prevent insider data exfiltration, refine security policy, and transform security effectiveness. Flying Cloud was founded in 2014. It is privately funded with offices in Polson, MT, Santa Cruz, CA, and Atlanta, GA.

Flying Cloud