SANS | Research Program

Survey

SANS Institute Survey: Do You Know Where Your Data Is?

Written by **Greg Scheidel** and **Terry Hicks**

April 2024

CrashPlan

Executive Summary

Endpoint data security is an increasingly serious concern for organizations and their security teams worldwide. One key reason is data sprawl: the generation, accumulation, and storage of amounts of data so enormous that the organization struggles to effectively implement data governance policies and practices. This can be especially difficult when data is stored and processed on user endpoints. These endpoints—any device by which a user accesses an organization's networks, applications, or data—take the form of an extraordinary and ever-expanding range of devices, from laptops to tablets to smartphones, and may be personally owned and, as such, less effectively managed than the organization's devices.

The ongoing shift to remote and hybrid work models means that these devices are frequently outside the physical control of the organization, and many organizations have visibility gaps for monitoring these devices off-network. This can lead to increased risk and decreased detection or response capabilities when malicious attackers compromise the device and access. New and sophisticated attack techniques are constantly being developed and executed that traditional endpoint security technologies, like antivirus and endpoint protection software, haven't kept up with. And open-source malware and other attack tools are making it easier than ever for even unsophisticated attackers to breach organizational defense. All these factors and more make endpoints an increasingly attractive attack surface for bad actors seeking to disclose, alter, and/or destroy an organization's data.

The growing awareness of these issues, and the risks they introduce, is the reason SANS Institute has chosen to conduct its first-ever survey of organizational practices in this area. SANS Endpoint Data Survey 2024 is designed to establish organizational stakeholders' views on the relative amount and makeup of data that exists on user endpoints versus central data stores, and the existence, rigor, and effectiveness of policies that either restrict or support storing data on user endpoints.

The survey results show that even though most organizations recognize the risks of storing sensitive data on endpoints, many continue to do so, even when they have policies and controls to limit the practice. This finding is especially concerning because a significant number of respondents report that much of the data they process and store are subject to rigorous regulatory compliance frameworks, including the European Union (EU) General Data Protection Regulation (GDPR), the U.S. Gramm-Leach-Bliley Act (GLBA), and the U.S. Health Insurance Portability and Accountability Act (HIPAA).

Interestingly, the risk most commonly reported as resulting from data damage or deletion was "operational," taking precedence over "financial," "compliance," "reputational," "strategic," or "legal" risk. The picture changes, however, when the issue is risk resulting from inappropriate data *exposure*. Then the most important risk is identified as "reputational," showing a clear concern for reputational impact even among non-management, security-focused respondents.

Key Findings

- **Organizations clearly recognize the risks of storing, processing, or transmitting sensitive data on endpoints, but many continue to do so, even when they have policies and technical controls in place to prevent the practice**
- **Regulatory compliance requirements represent a critical—and likely underrecognized—driver of the need for improved endpoint data security.**
- **The most common form of risk the survey respondents reported from endpoint security failures is operational. Reputational risk is, however, viewed as the most common form resulting from data exposure.**

Closing the identified gaps in endpoint data security—and, crucially, reducing the associated risks—will require organizations to first understand the reason for those gaps. Organizations might be tempted to implement or update policies or controls relating to data use on endpoints, and this may be a necessary step—but such policies and controls may be ineffective or even counterproductive if they don't align with the business needs and risk tolerance of the organization, and with the operational needs of the organization's users. Respondents with policies and controls to prevent users copying data to endpoints also report that their users do, in fact, copy data to endpoints. Organizations must consider how to manage—in an operationally sustainable and user-friendly way—data that will, perhaps inevitably, be stored on endpoints.

Survey Methodology and Demographics

This survey was conducted online, and collected self-assessed insights from 169 respondents in the US, Europe, and Asia, with the three largest industry sectors technology (20% of respondents), cybersecurity (13%), and banking and finance (11%). Nearly half of the respondents represented small or midsize organizations: 25% with 100 employees or fewer, 20% with 101–500 employees. Most the organizations (56%) have their headquarters in the U.S., while 12% are headquartered in Europe and 10% in Asia. A larger majority (61%) have operations in the U.S., 29% in Europe, and 25% in Asia, and 63% of the respondents directly support U.S. operations, 23% European operations, and 19% operations in Asia. In terms of the respondents' roles within their organizations, 30% reported acting in non-management security-focused roles, 17% in management security-focused roles, and 24% in management roles that are not security-focused. Figure 1 provides a snapshot of survey demographics.

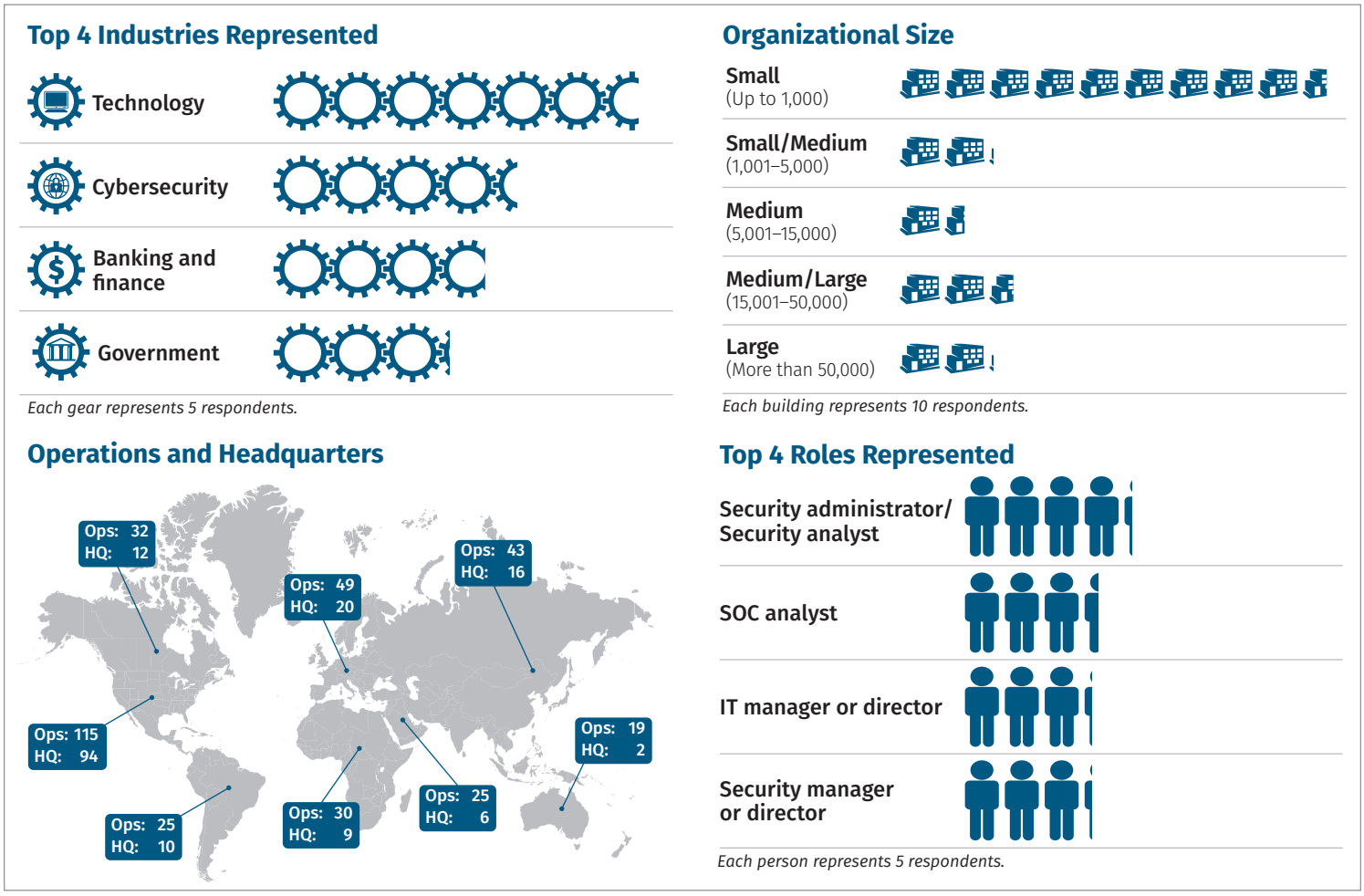

Figure 1. Survey Demographics

Endpoint Security: Organizations at Risk

One of the most notable survey results is the broad recognition that regulatory compliance requirements are—and should be—the most important drivers of the adoption of enhanced endpoint protection processes and technologies. A clear majority of the respondents report that their organizations store, process, or transmit data that requires special handling because of externally imposed regulations or other standards. PII data is handled by 79% of the respondent organizations, 65% handle technical and risk management data, 56% handle financial data, and 52% handle intellectual property (IP).

The GDPR and UK privacy frameworks are the regulatory standards that are most commonly reported as applying to all data types: 50% of PII, 46% of technical and risk data, 46% of financial data, and 53% of IP data. The Payment Card Industry Data Security Standard (PCI DSS) was a close second, applying to 49% of PII, 39% of technical and risk management data, 46% of financial data, and 38% of IP data. HIPAA applied to between 22% and 44% of those data types, and the U.S. Family Educational Rights and Privacy Act (FERPA) to between 20% and 24%.

While the respondents' organizations clearly recognize the risks of storing sensitive data on endpoints, and a majority of the organizations have data protection and retention policies in place, the survey results indicate that many continue to do so at levels inconsistent with these policies:

- 62% prohibit storing PII data on endpoints, and 56% have technical controls in place to enforce this, but 55% report some PII data stored in files on endpoints (and 19% report that up to 20% of their PII data is in files on endpoints).
- 57% prohibit storing IP data on endpoints, and 54% have technical controls in place for enforcement, but 57% report some IP data stored in files on endpoints (and 22% report that up to 20% of their IP data is in files on endpoints).
- 56% prohibit storing financial data on endpoints, and 57% have technical controls in place for enforcement, but 53% report some financial data stored in files on endpoints (and 17% report up to 20% of their financial data being stored in files on endpoints).

It's possible to speculate as to the reasons for these gaps between recognized need and real-world practice. Some organizations may have budget or other resource constraints, some may be heavily focused on other security issues, and some may face cultural resistance to changes in security practices. It is also possible that there is a disconnect between business needs to work at the endpoint with the desire to secure data centrally, which would further indicate the need to both secure and backup data at the endpoint.

The survey results also show interesting perceptions of the risks from data security failures. Operational risk is the type most commonly cited by the respondents as a driver of concern about endpoint data security, exceeded only by financial risk in the case of financial data and reputational risk in the case of inappropriate exposure across all data types.

The Survey Findings: A Closer Look

The key overall conclusion to be drawn from the Endpoint Data Security Survey is that respondents recognize the endpoint as a threat vector, and the risks of storing and processing data on endpoints, but don't believe their organizations are doing enough to address these issues. They identify clear gaps between their organizations' need to protect sensitive data of all kinds from endpoint security failures—as reflected in established policies and installed tools—and their real-world practices.

Now let's take an in-depth look at some of the most important survey findings, with a view to identifying the respondents' key endpoint data security concerns, their perceptions of their key risks in this area, and their views of potential gaps in their capabilities to address those risks.

Regulatory Compliance Is a Critical Issue Worldwide

Regulatory compliance is—and should be—a key driver of endpoint security. A clear majority (at least 78%) of the respondents report that their organizations store, process, or transmit some type of data that requires special handling because of externally imposed regulations or other standards. In particular, they report that their organizations are subject to region-specific regulations and standards, even when the organization is not headquartered in or does not have operations in a given region. This makes sense given that many companies serve customers across the globe and Internet, business travel, and remote/hybrid work models mean that employees can access data from anywhere in the world. These dynamics introduce additional layers of regional and country specific regulatory requirements and standards based on customer and employee location.

What Kinds of Data Are Organizations Handling?

- **Personally identifiable information (PII): 79.3% of respondents' organizations**
- **Technical/risk management data: 64.5%**
- **Financial data: 55.6%**
- **Intellectual property (IP): 52.1%**

Two European privacy frameworks—the GDPR and the UK Data Protection Act (DPA) of 2018—are the regulatory standards that are most commonly reported as applying to all data types: 50% of PII, 46% of technical and risk data, 46% of financial data, and 53% of IP data. The Payment Card Industry Data Security Standard (PCI DSS) was a close second: It applies to 49% of PII, 39% of technical and risk management data, 46% of financial data, and 38% of IP data. HIPAA applied to between 22% and 44% of those data types, and the U.S. Family Educational Rights and Privacy Act (FERPA) to between 20% and 24%. See Table 1.

Table 1. Frameworks Applying to All Data Types

	PII	IP	Fin	Tech
GDPR/UK privacy regulations	50.4%	52.5%	45.9%	45.6%
PCI	48.9%	37.5%	45.9%	38.9%
HIPAA	43.5%	30.0%	22.4%	30.0%
FERPA	20.6%	21.3%	23.5%	20.0%
GLBA	13.0%	21.3%	12.9%	13.3%
Other	17.6%	21.3%	22.4%	31.1%
FedRAMP	14.5%	18.8%	10.6%	14.4%

When, for example, we drill down to look at one of the most sensitive forms of data, PII, we see, that a slight majority of organizations (50%) are subject to GDPR/UK rules, even though only 12% are headquartered in Europe and 29% have operations there.

Similarly, non-U.S.-based organizations—including the 12% that have US operations—are subject to some U.S. regulations and standards, such as GLBA and HIPAA. And PCI DSS requirements of course apply around the world to any business that accepts, handles, stores, or transmits PCI data.

Organizations clearly recognize the risks of storing all these types of sensitive data, and an overwhelming majority of the respondents' organizations have data protection and retention policies in place, but the survey results indicate that many within their organizations continue to process and store data on their endpoints at levels inconsistent with these policies.

Operational Risk Is a Key Concern

The risk most commonly reported as resulting from data damage or deletion was 'operational'—higher than financial, compliance, reputational, strategic, or legal risk for all data types except financial, and a close second even for financial data. Operational risk and impact can, of course, take many forms, including system downtime and lost productivity of the person at the endpoint, the personnel required for incident response, the time and resources spent in locating, restoring, or recreating lost or damaged data, and the involvement of teams (for example, legal, compliance and human resources) outside security. See Figures 2 and 3.

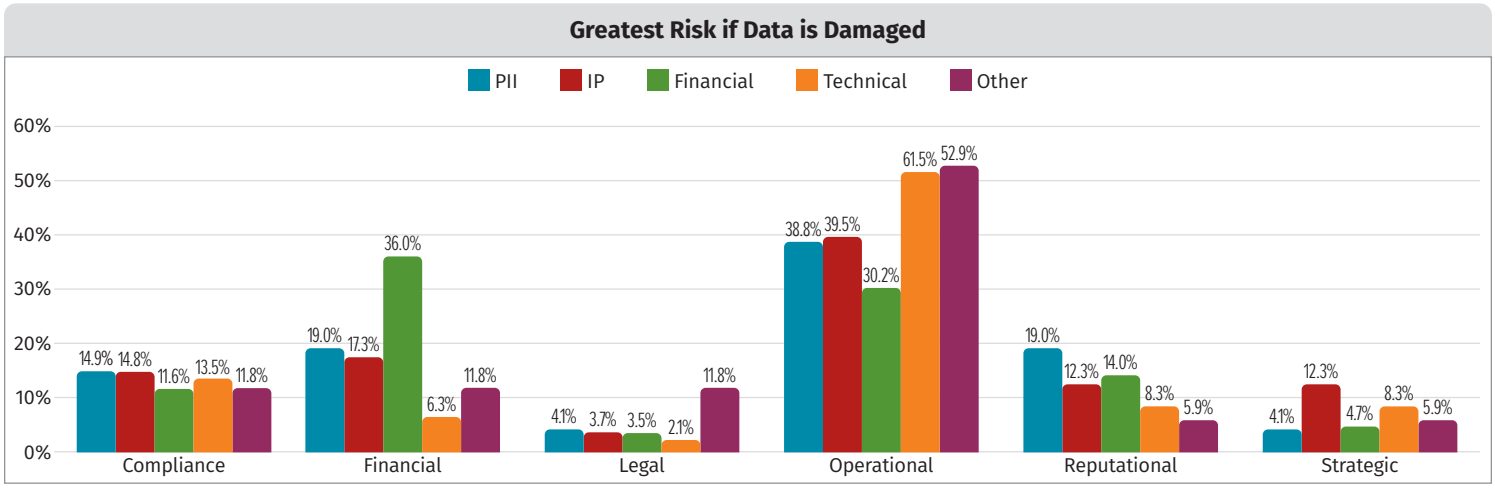

Figure 2. Greatest Risk if Data is Damaged

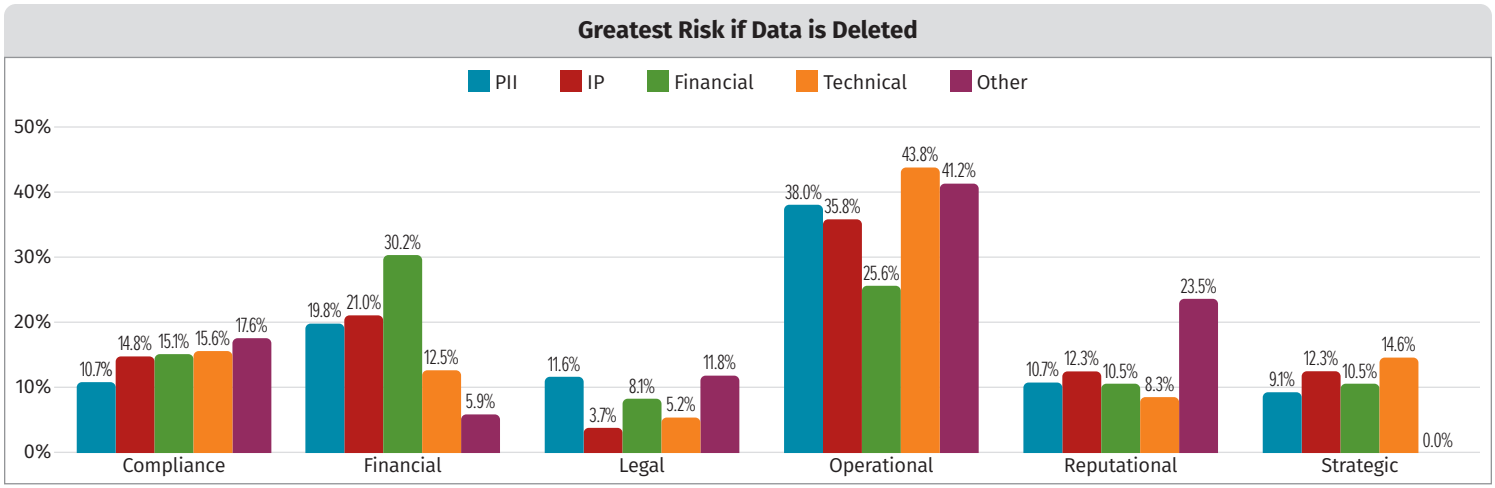

Figure 3. Greatest Risk if Data is Deleted

For business efficiency and especially the productivity of end users, operational risk is a more clear and present danger in their day-to-day work. While financial, compliance, reputational, strategic, or legal risks are all key to the organization, keeping their teams productive and efficient is also key. Data damage and loss at the end point is clearly seen as key to that.

The concern about operational risk is highest with regard to technical data, defined in the survey as "technical and risk management documentation such as network architecture drawings, system configuration backups, and vulnerability scan results." This is perhaps unsurprising given that 47% of respondents work in security-focused roles, and only a third of those respondents are in management-level roles. Concern about operational impact is understandable when the potentially damaged or deleted data is the very information needed to maintain and restore technical services!

Reputational Risk Is High for Exposed Data

The picture is strikingly different—and strikingly consistent—for cases of data exposure, as opposed to damage or deletion. The survey respondents reported that reputational risk was the highest concern for inappropriate exposure, and this was true across all data types. They were particularly concerned about exposure of the "other" data type, a catchall term for data that doesn't fit cleanly in the other data type definitions. This is understandable when one considers the real-world costs—which may bleed over into all the other types of damage—associated with damage to an organization's brand. See Figure 4.

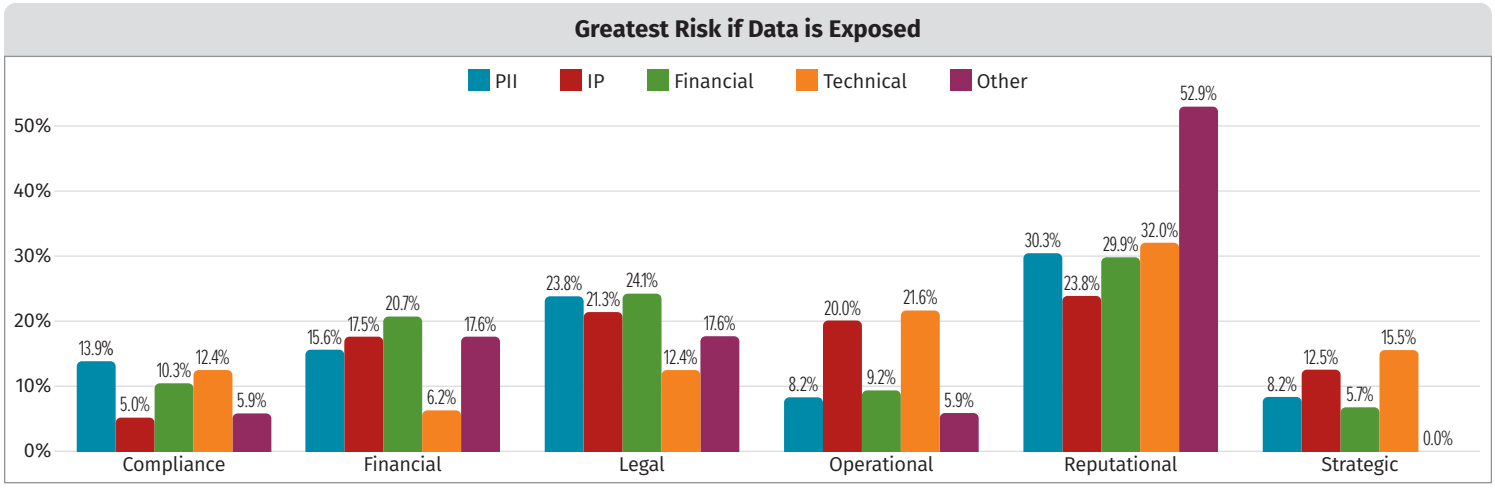

Figure 4. Greatest Risk if Data is Exposed

Where Organizations Are Storing Data

The survey results show that the respondents' organizations recognize the risks of storing sensitive data of all kinds on endpoints but are not satisfied with their current practices for protecting that data. Let's take a look at what the data shows for each type of data.

PII

A majority of respondents' organizations (56%) have technical controls in place to prevent storing PII on endpoints, and even more (62%) have policies prohibiting the practice—and yet a clear majority of their organizations' endpoints do in fact store at least some PII. The obvious conclusion is that their controls and policies aren't working, at least not to the desired level, possibly because of a disconnect between business needs and security or compliance needs. This suggests changes—with end user input—may be warranted in both policy and approach. See Table 2.

Table 2. Controls, Policies and Practices for PII Storage on Endpoints

Technical Controls Prevent Copying to Endpoints	56.4%
Policies Prohibit Copying to Endpoints	62.1%
Some Data Stored on Endpoints	**54.8%**
Some Data Saved on Endpoints	**79.7%**
Between 1% and 20% of Data Saved on Endpoints	32.3%
More Than 20% of Data Saved on Endpoints	47.4%

IP Data

Clear majorities of the respondents also report that their organizations have installed technical controls to prevent storing IP on endpoints and have policies in place forbidding it. And yet the reality is that substantial amounts of IP are in fact stored in this way. See Table 3.

Table 3. Controls, Policies, and Practices for IP Data Storage on Endpoints

Technical Controls Prevent Copying to Endpoints	54.0%
Policies Prohibit Copying to Endpoints	52.0%
Some Data Stored on Endpoints	**56.8%**
Some Data Saved on Endpoints	**78.2%**
Between 1% and 20% of Data Saved on Endpoints	32.2%
More Than 20% of Data Saved on Endpoints	46.0%

Financial Data

This trend holds true for financial data, as well, with majorities reporting controls and policies to prevent endpoint storage of this highly sensitive data, yet also reporting that this data continues to be stored in this way. See Table 4.

Table 4. Controls, Policies, and Practices for Financial Data Storage on Endpoints

Technical Controls Prevent Copying to Endpoints	57.4%
Policies Prohibit Copying to Endpoints	56.4%
Some Data Stored on Endpoints	**53.2%**
Some Data Saved on Endpoints	**76.3%**
Between 1% and 20% of Data Saved on Endpoints	34.4%
More Than 20% of Data Saved on Endpoints	41.9%

Technical Data

The picture is somewhat different for technical data, with somewhat fewer respondents reporting technical controls or policies to prevent endpoint storage, suggesting less concern about the damage, deletion, or exposure of this data. See Table 5.

Table 5. Controls, Policies, and Practices for Technical Data Storage on Endpoints

Technical Controls Prevent Copying to Endpoints	46.1%
Policies Prohibit Copying to Endpoints	48.1%
Some Data Stored on Endpoints	**46.8%**
Some Data Saved on Endpoints	**79.8%**
Between 1% and 20% of Data Saved on Endpoints	27.3%
More Than 20% of Data Saved on Endpoints	51.9%

Other Data

Policies and controls are even less common for the "other" data category. See Table 6.

Table 6. Controls, Policies, and Practices for Other Data

Technical Controls Prevent Copying to Endpoints	10.5%
Policies Prohibit Copying to Endpoints	26.3%
Some Data Stored on Endpoints	**30.0%**
Some Data Saved on Endpoints	**73.7%**
Between 1% and 20% of Data Saved on Endpoints	36.8%
More Than 20% of Data Saved on Endpoints	36.8%

File Storage

The survey respondents were asked to detail how much of their sensitive data is stored in files—as opposed to formal data systems—because this practice carries potential endpoint security risks. The key reason: Data in file format is more portable, and therefore more likely to find its way to endpoints. And the survey results show that large majorities of the respondents' organizations do in fact store at least some data in this way. See Table 7.

Table 7. File Storage Practices by Data Type

Some PII Stored as Files	74.0%
Some IP Data Stored as Files	77.9%
Some Financial Data Stored as Files	66.7%
Some Technical Data Stored as Files	72.8%
Some Other Data Stored as Files	63.2%

Additional Data Management Requirements

Survey data also confirmed respondent organizations' use of other data management requirements—implying the need for solutions that help meet those requirements. For example, significant percentages of the respondents report that their organizations have established specific restrictions for accessing different types of data under specific circumstances, such as data that only IT or legal are allowed to access under specific conditions, or data under legal hold. These restrictions break down as follows:

- Restrictions on PII storage: 43%
- Restrictions on IP data storage: 37%
- Restrictions on financial data storage: 39%
- Restrictions on technical data storage: 28%
- Restrictions on other data storage: 26%

Another area where the survey responses reveal organizational priorities is data retention period requirements:

- 79% of respondents have PII retention period requirements
- 74% have IP retention period requirements
- 78% have financial data retention period requirements
- 19% have technical data retention period requirements
- 58% have other data retention period requirements

All these requirements extend to these data types stored on endpoints (at least if the data isn't stored elsewhere).

Conclusion

The survey results show clearly that the respondents—and not only those whose functions are security-focused—recognize the importance of data security at the endpoint. Just as importantly, they broadly recognize that their current endpoint data security practices are inadequate, and in fact don't even meet their own established policies. Despite the widespread presence of policies prohibiting the storage of sensitive data of all types at the endpoint, and the installation of technical controls to prevent, enormous amounts of the most sensitive data is still found in these extremely vulnerable locations.

Organizations must revisit their approach to securing and protecting data at the endpoint, taking into consideration what business needs are driving users to work on their local devices while at the same time meet the security and compliance needs. The gaps between the risk to PII, IP, financial, and other types of data and the reality of day-to-day practices must be closed. Taking a zero-sum approach of prohibiting data at the endpoint has clearly not worked to-date and not from lack of effort of skilled security practitioners and management. So, more holistic approaches for endpoint data resiliency must be considered.

Sponsor

SANS would like to thank this paper's sponsor: