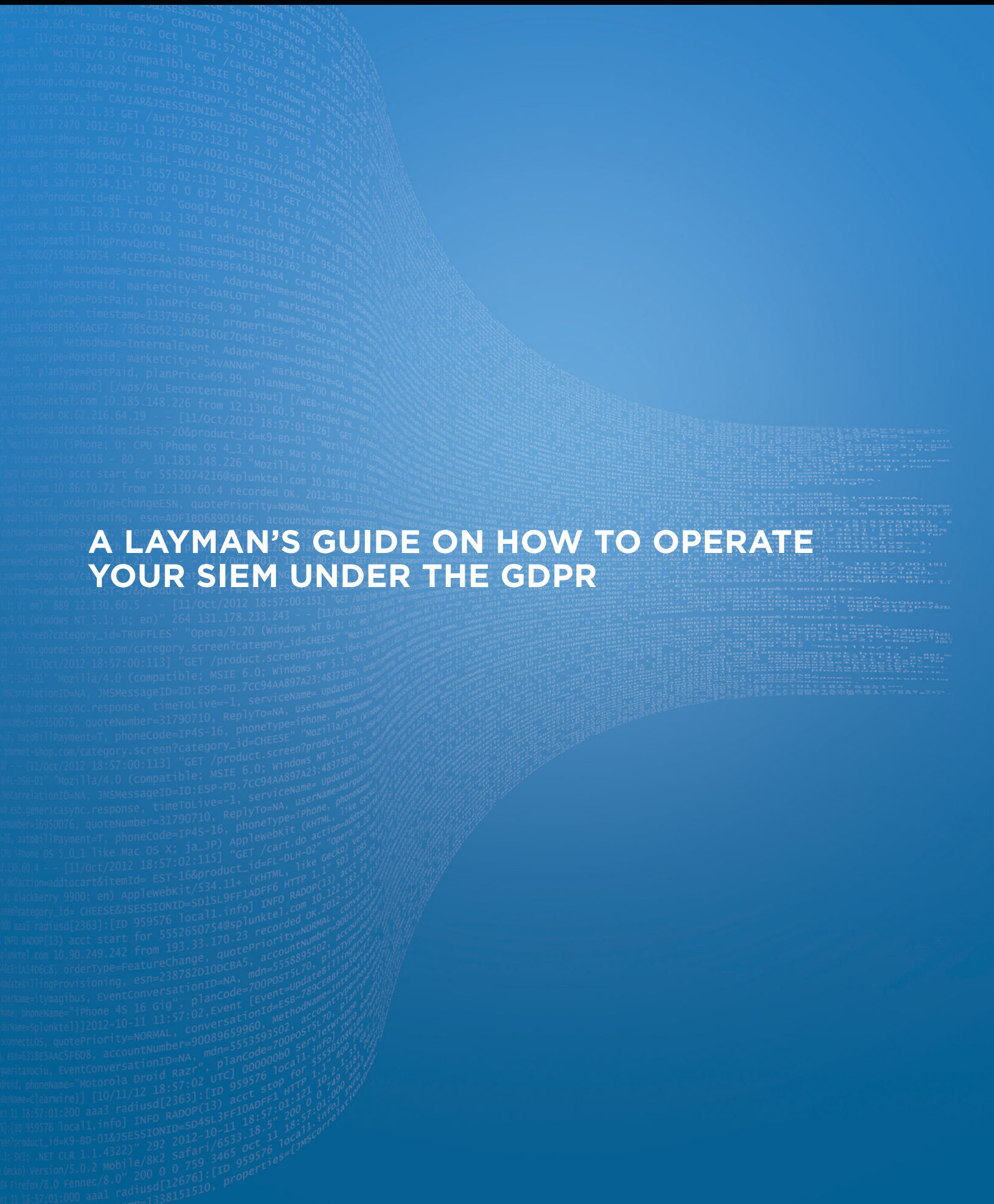
splunk>

WHITE PAPER

A LAYMAN'S GUIDE ON HOW TO OPERATE YOUR SIEM UNDER THE GDPR



Splunk invited Freddy Dezeure, former head of CERT-EU, to provide advice on how to use Splunk as a SIEM in compliance with the European Union General Data Protection Regulation (GDPR).

INTRODUCTION

The EU's General Data Protection Regulation, Regulation (EU) 2016/679, or "GDPR", takes effect on 25 May 2018 without the need for EU Member States to enact it into local law. The full text of the GDPR can be found here.

This Guide provides an overview of portions of the GDPR most relevant to processing log data using Splunk. The Guide has two parts:

- **PART I** provides a general introduction to the GDPR, highlighting aspects that are the most relevant to understanding the impact of the GDPR on log management.
- **PART II** provides specific compliance guidance and use cases for network and information security logs.

PART I: GENERAL INTRODUCTION TO THE GDPR

Scope of the GDPR

The GDPR covers the processing of personal data, broadly defined to include *any information relating to an identified or identifiable natural person*, including such things as telephone numbers, email addresses, IP addresses, MAC addresses, cookies, RFIDs, credit cards, geolocation data, etc., if it identifies a natural person directly or in combination with other information.

The GDPR's reach is expansive: it extends to organizations regardless of whether or not they are located in the EU or their processing is taking place in the EU. It also applies to the processing of personal data, even if it is done for free. This means that the GDPR impacts a large number of commercial and public organizations across the globe.

The GDPR applies to "controllers" and "processors" of personal data. A controller is the natural or legal person who determines the purposes and the means of processing personal data. A processor is a natural or legal person which processes personal data on behalf of a controller.

About Freddy Dezeure:

Freddy Dezeure graduated from the KUL in Leuven, Belgium, with a Master of Science in Engineering in 1982. He was CIO of ETAP NV from 1982 until 1987. He joined the European Commission in 1987 where he held a variety of management positions in administrative, financial and operational areas, in particular in information technology. He set up the EU Computer Emergency and Response Team (CERT-EU) for the EU institutions, agencies and bodies in 2011 and made it into one of the most mature and respected CERTs in Europe. Until May 2017 he held the position of the Head of CERT-EU. Presently, he is an independent management consultant providing strategic advice in cyber security and cyber risk management and acting as Board Member and Advisory Board Member in several high-tech companies.

A controller is responsible for implementing appropriate and effective measures to comply with GDPR and must be able to demonstrate its compliance. Controllers are responsible for verifying that processors acting on their behalf similarly comply with the Regulation.

Controllers and processors must maintain a record of their processing activities, documenting the kind of data being processed, the purpose of the processing, the parties with whom the data is shared, the data retention limits for the processed data, and the security measures taken to protect the data.

Article 37 provides that the controller and the processor shall designate a data protection officer (DPO) in cases where processing is taking place on a large scale or involves sensitive categories of data, such as health data or criminal records. The DPO oversees compliance with GDPR. The DPO should be properly resourced and report directly to the highest management of the controller or processor. The role of a DPO can be combined with other tasks and duties (like CISO or CSO) if they do not result in conflicts of interests.

Some general principles

The GDPR is guided by fundamental principles that personal data should be:

- Processed lawfully, fairly and transparently
- Collected for specific, explicit and legitimate purposes and not further processed in a manner incompatible with those purposes
- Processed only in so far as is necessary for the purpose of the processing
- Accurate and not kept longer than necessary for the purpose for which it is processed
- Processed in manner that ensures appropriate security and confidentiality (Article 5)

Personal data can be lawfully processed under the GDPR if it is done with the data subject's consent, which must be explicitly given and based upon clear and plain language regarding the purpose of the processing activity or based on other lawful grounds for the processing, such as the performance of a contract, compliance with a legal obligation, the protection of the vital interests of a natural person, the performance of a task carried out in the public interest or legitimate interests of a controller/processor (Article 6).

Data subjects have the right to access their data and to rectify, transfer and ask that it be removed when the data subject's consent is withdrawn. This is known as the "right to be forgotten".

Consequences of non-compliance with the regulation

A personal data breach is defined as a security incident leading to destruction, loss, alteration, unauthorized disclosure or access to personal data. Such breaches may lead to obligations of notification and to sanctions and penalties.

A controller should notify the supervisory authority of a personal data breach without delay and, where feasible, not later than 72 hours after having become aware of it **if** the data breach is likely to result in a risk to the rights and freedoms of natural persons. The risk to the right and freedoms can be minimized or avoided by deploying techniques such as pseudonymization and encryption.

Reference: Recital 85 and Article 33

The controller should also communicate to the data subject a personal data breach, without undue delay, where that personal data breach is likely to result in a high risk to the rights and freedoms of natural persons. This communication should be coordinated with the supervisory authority, which will carefully examine the technical and organizational measures you have in place to secure the data.

Reference: Recital 86 and Article 34

Non-compliance with the Regulation can in itself lead to sanctions and penalties following a claim by a data subject with a supervisory body if he/she considers that his/her rights are infringed. Recital 146 indicates that the controller or processor "should compensate any damage a person may suffer" from processing that infringes the Regulation. Infringements shall also be subject to administrative fines up to 20million€ or 4 percent of the worldwide annual turnover to be determined on a case by case basis, depending on the specific nature of the infringement as detailed in Article 83. Member States are also able to lay down rules on criminal penalties for infringements as set forth in Article 84.

The consequences of non-compliance with the Regulation can therefore be a combination of damage compensation to the victims, administrative fines and criminal prosecution.

PART II: WHAT'S THE RISK OF PROCESSING AND STORING LOG DATA IN THE CONTEXT OF GDPR?

Do network and information security log files contain personal data?

While CSIRTs/CERTs/SOCs might be able to detect bad behaviour via pseudonymized information or via automated behaviour correlations - at a specific point where a detection becomes an incident you will need to know which technical user account might have been involved to further investigate and mitigate the incident. This means in practice that machine data related to users and their behaviour recorded in network and information security log files contains personal data.

What does this mean for compliance with GDPR? Does it limit the usefulness of Splunk and log monitoring in general?

As set forth above, the GDPR sets out in Article 6 lawful grounds for the processing of data. The most relevant to this Guide is the "legitimate interest" basis, which provides that processing may be performed if necessary for the purpose of carrying out the "legitimate interests" of the controller and where it does not outweigh the interests or "fundamental rights and freedoms" of the data subject.

Recital 49 provides explanatory comments which help to interpret the "legitimate interest" basis and which clarify that processing personal data "to the extent strictly necessary and proportionate" for "ensuring network and information security" constitutes a "legitimate interest" under the Regulation.

It follows that consent from the data subject is not needed for log management carried out for the purpose of ensuring network and information security where the processing is necessary and proportionate.

Obligations in terms of documentation and notification

Under Article 30, you are required to document the purpose and extent of the processing (including how you meet the necessity and proportionality standards). In documenting the necessity and proportionality of the processing of data in network and information security logs, an organization should consider the severity and the impact of incidents that are likely to be mitigated by log recording, management and correlation.

Compromised computers are a threat to the privacy and security of users, customers, their organizations and others. Depending on the case, consequences could range from exposing data to disruption or loss of assets. Security analytics with Splunk relies on log files to help mitigate these risks and quickly detect incidents and remediate them before they harm or jeopardize your IT assets.

The GDPR requires that personal data be stored for no longer than necessary to achieve the intended purpose. In the case of network and information security, that may vary depending on the length of time needed to detect, scope or remediate an incident or other legal or regulatory record retention requirements. Therefore, consideration should be given to how long log files should be maintained to give you the necessary audit trail security investigations require, since many vulnerabilities go undetected for long periods of time. You will also have to "look back" in your records to understand the scope of the risk created by the incident.

Precautions to take to comply and maintain visibility

Preliminary analysis of logs, correlation and triage is increasingly automated. Human analysts only intervene when there is a need for human triage and assessment, and typically, this involves only a small subset of your security team. Therefore, the risk of handling personal data contained in the log files may be considered low.

Depending on the nature of the logs, the potential risk of exposure may differ. Some things to consider:

- **Netflow, DNS and legacy firewall logs** are frequently pseudonymized, but by combining information from different sources, links leading to an identifiable person may nonetheless be made. Accordingly, you may want to limit access to these links in combination to those investigating confirmed security incidents;

- **Host logs (Applocker, AV, host firewall) and Active Directory logs** could also be pseudonymized, but from a risk mitigation point of view, the risk of not doing so may be low or acceptable because on balance these logs do not contain much in the way of personal data;
- **Proxy, next generation firewalls and application logs** usually contain user names so that individual behaviour may be monitored where legitimately needed and proportional to the task, but these personal identifiers could be separated from the logs or pseudonymized or their examination deferred after initial triage of other log sources has indicated an increased risk of a security incident;
- **Email logs** contain references to individual users and their communications. Access to the entries in these logs could be limited to those investigating confirmed security incidents.

During an incident, the handling of personal data of the impacted users will need to be exposed to the security team. In such cases, both the individual user and the organization have an interest in the problem being resolved.

However, the exposure of personal data can be minimized and the risks reduced by limiting access to the members of the security response team involved in the handling of the specific incident. If you are processing logs that contain sensitive data sets under Article 35 (Data Protection Impact Assessment) the limited pool of people with access to the data can be used to help demonstrate the "proportionality" of the processing operations.

Risk of log file data breaches

Appropriate security measures should be put in place to mitigate the risk of breaches of the personal data stored in the logs. These measures should take into account the level of security appropriate to the risk of disclosure, modification or loss of the data.

When balancing the appropriate security measures against the potential risks, it's helpful to keep in mind the following:

- When considering the risk of unauthorized exposure of log files to third parties outside an organization, it's important to note that log files present no new or exceptional risks and should be treated much the same way the organization manages other security risks associated with its virtual assets.
- When considering the risk of unauthorized access within your organization, just as with other applications, networks and systems in your IT ecosystem, role based monitoring can be deployed to help prevent the use of the personal data stored in logs from being used for unauthorized purposes. Likewise, organizations should take appropriate measures (organizational, security policies and segmentation) in terms of access to the logs that help ensure that secondary uses of the personal data contained in them are prevented or conducted only with knowledge and prior review and approval.
- Limit the risk of unauthorized or unnecessary use of the personal data in log files by allowing only select individuals with your organization who have a pre-defined "need to know" to access them, such as the security team investigating security incidents.

In most cases the security measures your security team has already put in place to protect your IT infrastructure will suffice to mitigate the risk of external breaches presented by the use of log files.

Pseudonymization and encryption options

If your risk assessment suggests the need for heightened security measures around log files to mitigate risk, there are other options to consider.

One of the measures the regulation highlights to mitigate risks is pseudonymization, a technique which, if properly implemented, can reduce the risk that data can be attributed to a specific person without additional information that is stored separately. In the case of network and security logs, this means splitting off certain data (usernames in proxy logs, recipient address in email logs) in processing the log files or in the access procedures and making them only accessible on demand (in case of a confirmed security incident). Encryption may also help, but of course, is of limited use if the legitimate credential access control is breached.

Splunk can facilitate the implementation of these options. Different techniques and their advantages and drawbacks are described here.

Specific risk reduction use cases

Leveraging machine data for security analytics with the Splunk platform helps support many key GDPR security requirements. Here are some examples:

1. Centralize log files and indexes and secure them

Log files and machine data generated by the organisation's IT infrastructure and applications are stored in Splunk in a way that allows you to secure and control access to the data within Splunk. By centralizing the log repository and organizing its protection, the organization helps fulfil the mandate of set out in Article 32 (Security of processing) of the GDPR to "implement appropriate technical and organizational measures to ensure a level of security appropriate to the risk".

Using Splunk, you can:

- Install the operating systems and software in a secure manner and harden its controls;
- Secure log forwarding by encrypting the communication with signed certificates and monitoring the proper functioning of the forwarders;
- Deploy role-based access control and two-factor authentication;
- Secure browser configurations and encrypt web communications;
- Monitor the use of Splunk and detect anomalies as you would with any other critical service.

2. Pseudonymize personal data in log files

As indicated in this guide, Splunk supports pseudonymization of personal data in log files at different layers to provide an additional level of protection where needed. Two possible pseudonymization methods are described here:

- By event duplication: sort log files into different indexes; one with pseudonyms accessible for detection and triage; one with the full data set accessible for incident response combined with access control procedures;
- By transforming the search index into a pseudonymized index and using the latter for detection and triage and the former for incident response.

Log management and correlation in support of GDPR

Splunk is a platform to store, manage and correlate machine data and can be used to support a comprehensive information security risk management system by fostering early detection and correlating such findings with key information to support data breach impact assessments. Three real world scenario's, what they mean under GDPR and how machine data helps with can be found in Splunk's white paper **"How Machine Data Supports GDPR Compliance"**.

The design and implementation of a system that provides early detection and data breach scoping by correlating events in log files is an "appropriate technical and organisational measure" designed to ensure a level of security appropriate to the risk, which is what the GDPR security standard is all about.

Want to learn more about operating your SIEM solution under the GDPR?

Listen to our webinar **"A Day in the Life of a GDPR Breach."** Or read our white paper **"How Machine Data Supports GDPR Compliance"** and discover how to be prepared come May 2018.

splunk>

Learn more: www.splunk.com/asksales

www.splunk.com