Perlego

Data Security Whitepaper For Publishers

Perlego Limited
26 Hatton Gardens
London EC1N 8BR
perlego.com

Table Of Contents

Introduction

This document outlines the security measures taken at Perlego to protect publisher content. This whitepaper is confidential and must not be shared with any unauthorised parties. This whitepaper gives an overview of Perlego's security practices throughout its organisation and its technology. This document also includes information on the control mechanisms in place to ensure that publisher content is accessed in a secure and controlled way. At Perlego, we understand the security concerns that publishers have regarding their content and how it is processed, stored and accessed on the Perlego platform. That is why security and the protection of data is a top priority that drives our organisational structure and technology.

Perlego's security culture

In Perlego's hiring process, Perlego will verify an individual's education and previous employment, complete with reference checks. The extent of further security checks such as criminal, credit, immigration and security checks is dependent upon the desired position.

All Perlego employees undergo basic security training as part of the orientation process. Depending on the positions, employees will receive ongoing and specialised security training throughout their career at Perlego.

Perlego employs certified security and privacy professionals who are part of our software engineering division who focus on information, application and network security.
Perlego works with Ametros Group to ensure full compliance with GDPR.

Technology & Operational security

Vulnerability Management

Perlego actively scans for security threats as part of its vulnerability management process using a combination of commercially available penetration testing tools, quality assurance

processes, third party package auditing tools, software security reviews and code quality analysis tools.

Perlego periodically scans its production environments for vulnerabilities using penetration testing tools and also regularly works with accredited third party penetration testing services to secure its platform. A combination of manual and automated penetration tests are regularly conducted across the application and its underlying infrastructure. Perlego has behavioural analytics tools deployed in the cloud that can monitor for and detect a multitude of security threats throughout the entire Perlego network.

Once a vulnerability requiring remediation has been identified, it is logged and prioritised before being assigned an owner. If a vulnerability is discovered within the production environment or a vulnerability is found to have an impact on the security of user data and/or publisher data, it is given the highest priority and is to be remediated as soon as possible. Depending on severity, a vulnerability will be escalated to the incident response team for remediation and to ensure that the vulnerability has not been exploited by any malicious parties through forensic analysis of system logs.

Malware prevention and monitoring

Perlego takes malware threats very seriously as malware can lead to account compromise, data theft and possibly unauthorised access to internal systems. This is why Perlego uses a variety of methods to prevent, detect and remove malware. Automated virus scanning (using various antivirus and antimalware tools) is installed on all employee machines and throughout Perlego's cloud infrastructure. Internal traffic and administrator activity is automatically inspected for suspicious behaviour using a variety of tools to capture and analyse network traffic logs, system logs and firewall logs, in combination with threat detection tools (Cloudflare, AWS GuardDuty and DarkTrace). These tools can reveal inbound and outbound communications to and from suspicious or even malicious third parties from malware that may have found its way into the Perlego network.

Incident management

Perlego has an incident response team that is available 24/7 to remediate security incidents that may affect the confidentiality, integrity, or availability of systems or data. Every member

of the team is a full time employee and is properly trained to respond to security incidents. Perlego has an Incident Response Plan (IRP) that it uses to remediate security incidents. It is part of the IRP to notify all affected parties following an incident within 48 hours.

Physical Infrastructure Security

Perlego makes use of Amazon Web Services (AWS) to host its entire infrastructure. In doing so, Perlego benefits from the state-of-the-art data centers managed by Amazon and the extremely high degree of security that is implemented throughout their IT infrastructure.

Read more about Amazon's cloud security and data controls:

- Cloud Security – Amazon Web Services (AWS) https://aws.amazon.com/security/
- AWS Data Centers – Amazon Web Services (AWS) https://aws.amazon.com/compliance/data-center/

Firewalls

Perlego implements strict firewall rules throughout its infrastructure. Using Amazon security groups and a defense in depth approach to security, the Perlego infrastructure is composed of many services that have strict access controls on outbound and inbound connections in order to limit the impact of a security incident. This helps prevent unauthorised threat actors moving laterally across the network.
Perlego also makes use of Web Application Firewalls (WAF), used for filtering, monitoring and blocking HTTP traffic to and from the platform. Perlego primarily makes use of Cloudflare's WAF.

Securing data in transit and at rest

Data is at risk when it travels across the internet or within networks. Perlego uses strong encryption protocols such as TLS to secure the connections between customer devices and Perlego's web services and APIs. In addition to this, all sensitive data within the Perlego infrastructure is encrypted at rest using the industry standard encryption schemes such as the Advanced Encryption Standard (AES) 256-bit encryption.

Data access and restrictions

Administrative access

Only a small group of employees have access to publisher content. Access rights and levels are based on job functions and roles. We employ the concepts of least-privilege and need-to-know to match access privileges to defined responsibilities. All access to publisher content is made on Perlego owned employee machines that are periodically scanned for vulnerabilities, fully encrypted and can be remotely wiped (in the event of loss or theft).

Third-party services

Perlego will not share publisher content with third parties without written permission to do so. Perlego does not work with any sub distributors.

Content Processing

Publishers are required to send their content over a Secure File Transfer Protocol (SFTP) connection where content is encrypted in transit. Content is uploaded to a secure AWS S3 bucket where content is encrypted at rest in what can be described as a Demilitarized Zone (DMZ).

All files are sanitised and scanned for viruses. The sanitised files are encrypted each with their own unique and cryptographically secure encryption key and stored in AWS S3 Glacier. Publisher content is then automatically moved into a serverless processing pipeline where content is transformed into a format that can be securely served to users of the Perlego platform. Each book in its processed form is also encrypted with its own unique and cryptographically secure encryption key using AES 256-bit encryption.

The location of the processed content that is served to users through the Perlego platform is in an entirely separate Virtual Private Cloud (VPC) to where the actual processing takes place. Network segregation plays a crucial role in securing publisher content.

eBook Security Access Controls

How users access publisher content

We offer a streamed on-demand delivery service via our web application. Our e-Reader is available through all modern browsers. Content is also accessed by our users through our mobile application. Content is downloadable but only through our mobile application and is only accessible within the application.

Use of DRM

Traditional eBook retailers offer eBooks for download that a user can open with a wide range of different eReaders. Digital Rights Management (DRM) software is used to protect the file that is downloaded and to control how the user can access the content in order to ensure it does not get copied and redistributed leading to a loss in sales and revenue for the publisher.

At Perlego, the use of traditional DRM software is not necessary, this is because eBook files are never made available for direct download. Content is instead served to users in two ways;

(1) From within the web application, content is streamed to the user in a controlled and secure way. Publisher content within the web application is accessed from a web browser and is protected with a series of strict copy, paste and print functions in addition to a set of service level security controls to prevent malicious behaviour (see below).

(2) From within the mobile application, it is possible to download the contents of an eBook but the user can only access and view the content through the mobile application. Strong encryption mechanisms are employed to ensure the files cannot be directly accessed by the user. The content is downloaded with two layers of encryption (AES 256-bit encryption and TLS). Each book is encrypted utilising a different key for each user and for each book.

Copy Controls

We take reasonable measures to prevent our users from copying content from our eReaders:

- We apply protection against bot scraping (using web application firewalls, recaptchas and behavioural analytics tools)
- We apply a maximum limit to the total number of books accessed
- We disable native copy/paste/print functionality in the browsers
- We limit copying to 10% of the book (unless agreed otherwise with publishers)
- We disable print functionality in the browsers
- We ensure that an eBook is never loaded into the user's browser in its entirety (by clearing the memory of old content as new content loads) – chapters and pages are progressively loaded and removed from the browser as the user reads a given book
- We ensure that eBook content is never stored in the user's browser cache

Service Security

Each user on Perlego is tracked via their IP address in order to obtain geolocation information. Users who are granted access to publisher content must first provide us with valid contact information and valid payment details. A user cannot create multiple accounts with the same card details, we fingerprint card details in order to impose this restriction using our third party payment provider, Stripe.

Geolocation "velocity" checks are in place, our service detects unreasonable changes in location (i.e. if a user logs into an account from different countries within a short period of time, access is denied). Geolocation is used to ensure that users can only access publisher content in specific geographies as per the agreed contractual limitations of the sale of content. For example, if a publisher only has the rights to sell a book in France, then only Perlego users who are in France can access that book.

There are limits in place on parallel connections; a user can have up to two devices concurrently logged in within the web application, and the user can have up to two devices concurrently logged in within the mobile application.

Our web and mobile applications limit the total number of books that can be opened or downloaded by a user within the space of 30 days. Our mobile application has strict access

controls to the content which can ensure that access expires after 30 days if their subscription is not renewed or if they disable access to the internet.

User Tracking

Perlego tracks key actions that a user takes from within both the web application and the mobile application (even when offline). This is to monitor how much of each eBook users have read. The percentage of a book read by a user factors into the royalty payments to publishers. A user's actions within the eReader are kept for audit purposes and can be used to demonstrate exactly how the royalties were distributed to the publishers (all personally identifiable user data and publisher data is kept anonymous in any such audit).