info@readspeaker.com
www.readspeaker.com

ReadSpeaker webReader and Storage of Personal Data

This document describes how the web application service ReadSpeaker webReader (webReader for short) handles personal data.

Details may change in future versions but will essentially remain the same.

How webReader Works

The content to be read aloud is either sent from the client web browser (POST) or fetched by our servers (GET). This depends on the implementation chosen. The data must be sent over an encrypted channel using HTTPS (TLS >=1.2).

The content is processed in the ReadSpeaker environment, converting the content to speech. When this is prepared, a redirect is sent back to the user which triggers the web browser to go to our servers that generated the audio.

No text content needs to be stored permanently for this to happen, data is only stored temporarily while the audio is being created.

Data Centers

ReadSpeaker has its own equipment at three data centers located in Stockholm/Sweden, Ashburn VA/USA, and Tokyo/Japan. The data centers mainly serve their respective regions. Additional data centers may be set up in the future.

It is possible to lock the webReader service so it is served exclusively from one of these data centers. Note however, that there is then no failover if that particular data center is unreachable, which means the service stops working.

We use several external CDN vendors to distribute all static files. See section CDN Vendors below.

Collected, Stored, and Used Data

End user data such as user agent, OS type, device type, cookies, referrer data, IP address, geolocalization is stored only in temporary web server logs.

If personal information is present on a website and is available to be read, it can be sent over to our data centers to be turned into sound files. If there is no personal data on the website, no data will be processed by webReader.

There are two kinds of logs that are used for different purposes.

Web Server Logs

These logs are necessary for troubleshooting and to prevent illegitimate usage. The log files are standard web server logs. The access to these logs is restricted to a limited number of people with specific roles related to maintenance. Web server log records older than 30 days are deleted permanently.

The log information is restricted to technical staff and only reachable through encrypted channels using MFA. ReadSpeaker does not sell or pass on any log information.

Statistics Logs

The data for statistics logs is anonymized directly. It only contains customer ID, timestamps, and which voice was used. This data is used for billing.

Statistics are saved per activation. An activation is when a user clicks the Listen button to get content read aloud.

The statistics log contains what URL the button was activated from. Unless the URLs are unique for each user, no user info is saved. If that should be the case, the implementer of the webReader service can override what should be stored so that the URL does not get stored.

External Party/Subcontractor

Translation Feature

The translation feature uses APIs from Google Cloud or DeepL to translate content. Which provider is used is configurable per customer.

All translation requests, regardless of provider, go through ReadSpeaker's servers. Only the content to be translated is transferred to the third-party API, never IP addresses or other personal information belonging to the end user.

If a customer does not want the translation feature to be available in webReader, it can be turned off in the customer's configuration.

More information on the handling of data by Google can be found on the Google Cloud Translation API help pages https://cloud.google.com/translate/data-usage.

More information on the handling of data by DeepL can be found on https://www.deepl.com/en/privacy.

Dictionary Feature

The dictionary feature uses APIs from Oxford Dictionaries, Wiktionary, or dictionaries running locally in the ReadSpeaker environment to look up definitions of words. Which provider is used depends on the language.

All dictionary requests, regardless of provider, go through ReadSpeaker's servers. Only the single word that is to be looked up is transferred to the third-party API, never IP addresses or other personal information belonging to the end user.

If a customer does not want the Dictionary feature to be available in webReader, it can be turned off in the customer's configuration.

CDN Vendors

For the distribution of static files such as JavaScript, CSS, media files, etc., we use external content delivery network (CDN) providers.

The distributed files are required for the functionality of the player. When a browser fetches files from the CDN the IP, the user agent, and possibly the referer are unavoidably sent to the CDN. However, the customer can always choose a CDN where traffic remains within the specified region. Nodes within regions are selected to meet the requirements of regional privacy regulations. ReadSpeaker makes sure to provide CDNs for both global storage and regional storage, depending on what control the customer wishes to have over where the information shall reside.

The distributed files do not contain any content from the customer website. Also, they cannot be used in any way to identify or gather any personal information and are therefore safe to distribute publicly over the Internet.

If a customer wants to have full control over how and from where static files are fetched, the customer may host these files themselves. Doing so will move the responsibility for updating the content to the customer because as a result, ReadSpeaker cannot perform automatic updates of webReader.

Caching

Data that is required for webReader to work optimally consists of the audio files, and these are by default only stored temporarily in our cache, on disc. Cache clearing is done individually for each file. A file's caching time is extended as long as it continues to receive requests. While the exact caching duration may be adjusted for capacity management purposes, cached content is never stored for more than 30 days. When webReader detects changes on a customer website, it will create a new cache entry to ensure the latest content is delivered.

It is possible to turn off the audio caching, for example if the content on the website is classified. This will affect the pricing due to the increased capacity consumed by not taking advantage of the caching system and for additional development needed.

Distribution to Third Parties

There is no distribution of collected data.

Cookies

webReader sets two cookies:

1. An initiation cookie that determines whether or not to load the scripts on page load. This cookie has the name "_rspkrLoadCore" and is a session-only cookie. The cookie is set after webReader has been activated, i.e. when there has been an interaction with the player. This cookie is used to let us know that a user has activated webReader on a page. We will then automatically load the ReadSpeaker scripts when the user navigates to a different page, so the user will get audio quicker.

2. A cookie to store changes of settings in the settings menu. By default, this cookie is called "ReadSpeakerSettings", but its name can be changed by configuring "general.cookieName". The cookie has a default lifetime of 360000000 ms (i.e. about 4 days) which can be changed with the configuration "general.cookieLifetime". This cookie is used to save individual users' settings, so that they do not have to re-apply their preferred settings on every page they navigate to.

Data Purposes: webReader saves the users' settings in a cookie, so they are remembered and applied when the user navigates between pages.

Data Collected: No data is collected. Information about speed, highlighting colors, etc. are saved in the cookies. No IP address is stored.

Data Recipients: No data from the cookie is sent anywhere, it's only stored in the browser.

Conditioning for Cookies

If a customer wishes to require consent from users who click the Listen button, webReader can be configured to display a message that requires the user's consent before playback is started.

It is possible to turn off the cookies. If this is done, users that prefer to use webReader with personalized settings such as reading speed, highlighting colors, etc., will have to change the settings from the settings menu every time they load a new page.

GDPR

Since May 2018, ReadSpeaker is fully compliant with the EU regulation - General Data Protection Regulation (GDPR).

Data Privacy Matters

For inquiries regarding privacy matters, please contact our data protection team through gdpr@readspeaker.com.

Further Questions

If you have further questions, please contact our support team via email, support@readspeaker.com.