

Lingsoft Speech Service - Data protection description

Effective as of August 18 2021.

Introduction

Lingsoft Speech Service processes audio data to the text format (Speech Recognition Data). Lingsoft Speech Service ensures that appropriate measures are in place for the protection of the data that are processed for users.

Confidentiality

Lingsoft Speech Service has comprehensive **Information Security Principles and Procedures** in place that must be followed by all persons working for the Lingsoft Speech Service. These Principles and Procedures define all important security aspects to keep the use of Lingsoft Speech Service secure.

All users need to authenticate in order to use Lingsoft Speech Service, users get access rights according to the roles they have in their organisation and their access to the data is restricted accordingly.

The operation of Lingsoft Speech Service requires that maintenance persons have access according to their work roles to the system where speech recognition data is processed and personal data is stored. These persons are prohibited from using their access permissions to access or view customer data unless it is necessary from the service point of view to do so.

Data to be protected

Users need to have a user id that is approved by Lingsoft Speech Service, either provided by Lingsoft Speech Service, or the user id of their own organisation when SSO authentication for the organisation is supported by Lingsoft Speech Service.

In order to be able to provide the right service for users, a user profile is created by Lingsoft Speech Service for users, which defines the details of which of Lingsoft Speech Service services the user can request and how the processing steps for the service requests are configured. User profile may also contain some personal data about the user. What personal

data is stored in the user profile is explained in the **Privacy Policy** of the Lingsoft Speech Service. User profiles that are in-activated by the Customers are removed from the system after one year of in-activity.

If the user is an external application then the user profile is made for that application, not for the end-users of that application. In that case an external application must provide the user id and organization name of the end user for each session that it does on behalf of the end user to the APIs of Lingsoft Speech Service, unless otherwise agreed separately. The use of Lingsoft Speech Service is identified by the user id of the caller and data processed by Lingsoft Speech Service is stored in the permanent storage (unless In-Memory processing configuration is on) of Lingsoft Speech Service for every speech recognition session that users make using the APIs of Lingsoft Speech Service. **Statistics** is used internally to monitor and further develop the service, and to invoice the users of the Lingsoft Speech Service. Statistics data is deleted after it is not anymore needed for reporting, billing or accounting purposes, normally statistics data is kept to the end of the next calendar year but it is deleted at the latest after ten years.

The data that is processed for the users by Lingsoft Speech Service are audio and video data called **speech recognition data**. Lingsoft Speech Service transcribes this data and returns it back to the users, who can correct it and if so desired, save the corrected transcription back to the service. In the core of the service are Speech Models that are responsible for transcribing the audio data. These Speech Models can be trained or improved with speech recognition data when users so request. In that case the speech recognition data that is produced during the processing is saved for training and used later in the training of the Speech Models. Speech recognition data that is stored in Speech Service is removed when it is no longer needed for training the Speech Models

Data processing options

Customers can request **in-memory processing** for speech recognition data, which means that speech recognition data is kept only in memory (not saved to permanent storage) and deleted right after the users end the use of speech recognition.

If in-memory processing is not used, users can request that speech recognition data is stored and used for **training of the speech model(s)** for their organisation. In that case the speech recognition data is put aside for the training purposes.

Data communication

Data communication from client applications (including Lingsoft Speech Service UI) to APIs of Lingsoft Speech Service are always secured using https-protocol in order to keep the traffic from

being monitored by third-party agents in the open network. Lingsoft Speech Service supports some of the most advanced commercially reasonable cipher suites to encrypt all data in transit. Cipher suite choices are upgraded as the landscape of the cryptographic changes.

Service infrastructure

Lingsoft Speech Service is hosted in data centres operated by Lingsoft or industry-leading service providers that offer state-of-the-art physical and other protection for the cloud infrastructure underlying Lingsoft Speech Service. These providers are responsible for restricting access to the above infrastructure to authorised personnel only.

Network access to Lingsoft Speech Service production environments from open, public networks is restricted. Only a small number of production servers are accessible from the internet. Only those network protocols essential for delivering the Service to its users are open at our perimeter.

Service Availability and Continuity

Lingsoft Speech Service is committed to provide highly-available service. Lingsoft Speech Service runs on servers in the data centers that are fault tolerant. Third-party software versions are updated continuously. Service capacity is automatically monitored and in case of lacking resources the corrective measures will be taken promptly. Lingsoft Speech Service is automatically monitored for failures and the operations team is alerted in case of a failure in order to quickly resolve unexpected incidents. Servicedesk of Lingsoft Speech Service processes incidents from the users according to the valid Support Agreement. Disaster-recovery plan and backup measures are in place.

Logging

Lingsoft Speech Service maintains a logging system for information concerning security, monitoring, availability, access and other metrics about the Lingsoft Speech Service. These logs are analysed for failure and security events. Log files are normally kept for 30 days but important log files may be kept longer.

Continuous improvements

Lingsoft Speech Service is continuously improved with updates and new features. The security aspects are regularly reviewed internally. Before releasing updates, they are pre-production tested in the testing environment.

Changes to data protection

Lingsoft may revise this description from time to time to reflect changes to the Lingsoft Speech Service, applicable laws, regulations or standards or other changes that may occur in the business. Lingsoft shall post the revised description on the same webpage where it published this description. Lingsoft may also use Lingsoft Speech Service, email or other means for notifying customers of such description changes. The revised description (or, as applicable, the new one) will be effective when posted as described unless the document itself specifies a later time for its entry into force.

Log of changes

Date	Short description
18.8.2021	Updated the section Availability, it is now titled "Service Availability and Continuity" and contains more information on these matters.