# SoapBox

# Our approach to data privacy & security

June 2023

# Contents

# Our approach to privacy

**SoapBox Labs is a privacy-first speech recognition company**. Since our founding in 2013, we've invested in a privacy-by-design approach to our data and technology, with the goal of protecting every kid's fundamental digital privacy rights. End-user voice data will never be sold to third parties, and will never be used to support activities like marketing and advertising.

> **It wasn't faster or cheaper to design in privacy, but as a deep-tech, kid-focused company, success for SoapBox meant respecting every kid's fundamental to privacy.**
>
> **Dr. Martyn Farrows, CEO, SoapBox Labs**

## Privacy by design

**Privacy-by-design means that we proactively take steps and build processes that ensure we never identify an individual child when processing their voice data**.

Throughout the provision of our services, we ensure that children maintain their fundamental digital privacy rights by minimizing the processing of personally identifiable information (PII) to the extent necessary to provide and enhance our services. (A copy of the Privacy Policy for our Online Speech Services is available on our website.)

We also ensure that

- All voice data sent to our system is anonymized.
- All voice data is encrypted in our system (in transit and at rest).
- No other data relating to the identity of the child (PII) is received by our system.

- Any voice data retained in our system is only used to improve the accuracy of our voice engine and is never used for marketing or shared with third parties for any other purpose.
- Our clients decide on the jurisdiction for data processing and whether data is retained post-processing for product improvement purposes. This decision is based on our client's relationship with the end user and their legal basis for processing the end-user voice data.
- Our system is designed to handle prompted utterances for literacy and language learning. In the very rare event that an utterance may contain spontaneous utterances of personal or sensitive data, we have strict protocols in place for immediate deletion, erasure, and notification (based on a Data Protection Impact Assessment, DPIA).

**SoapBox does not take any steps to re-identify or de-anonymize any voice data** or otherwise associate the voice data with personally identifiable information. We also do not authorize, instruct, or encourage any third party to do so.

## Regulatory compliance

Safeguarding children's online privacy is of primary importance to SoapBox Labs. We are certified under the PRIVO Kids Privacy Assured Program to ensure we are fully compliant with relevant data privacy and data protection regulations, including COPPA and GDPR. We are certified for PRIVO'S GDPR*kids* Privacy Assured Shield and COPPA Safe Harbor Certification annually.

**SoapBox has been committed to a privacy-by-design approach to building speech recognition technology for kids since 2013. PRIVO is proud to call them a customer and a partner.**
**Denise G. Tayloe, Co-Founder & CEO, PRIVO**

# Data security

SoapBox provides speech technology via cloud-based RESTful Web Services deployed on Microsoft Azure. Our application is deployed in multiple data centers globally and uses latency-based routing to route user requests to the data center closest to them. Each data center contains Load Balancers that handle incoming requests and securely routes them to underlying backend servers that handle each request. All communications with SoapBox Web Services are secure and encrypted using HTTPS & SSL.

SSL (Secure Sockets Layer) is a standard security technology to create an encrypted link between a server and a client. This link ensures that all data passed between the server and the client remain private and secure. It was designed to support protocols such as FTP, HTTP, and TELNET.

Hypertext Transfer Protocol Secure (HTTPS), or "HTTP Secure," is an application-specific implementation that is a combination of the Hypertext Transfer Protocol (HTTP) with the SSL/TLS. HTTPS is used to provide encrypted communication and secure identification of a server, so that no middle man can easily intercept the data.

## Does SoapBox process PII?

Other than the end-user voice data, SoapBox does not require any personally identifiable information (PII) for our voice engine to perform, and we only store the minimum amount of request metadata, including the user's device type and the country and region the request originated from. All data is encrypted in transit and at rest.

Once the request has been processed, we offer two options depending on the customer's requirement.

1. We store the audio file in our Cloud Blob storage, and a record of the file in our database. All data stored in Cloud Storage is within private buckets (not publicly accessible) and is encrypted at rest. Access to this data is restricted within SoapBox as per our privacy policies.
2. We delete data immediately after processing and do not store the audio file in our Cloud Blob storage. In this scenario, we store a record of the request, but not the actual audio recording.

To discuss our data privacy and security policies and protocols in more depth, please get in touch with us directly.

# Privacy resources

As a company committed to protecting kids' data privacy, we also advocate for other voice technology and kid-focused companies to do the same.

Here are some of our popular privacy resources and publications:

- *Voice Prints and Children's Rights* - Submission to the UN Office of the High Commission for Human Rights
- *Let's Talk Voice Tech, Data Privacy, and Kids* - CEO Dr. Martyn Farrows' guest post in Voicebot.ai
- *Data Privacy Matters for Kids. Here's Why* - Interview with children's data privacy expert Dr. Veronica Barassi
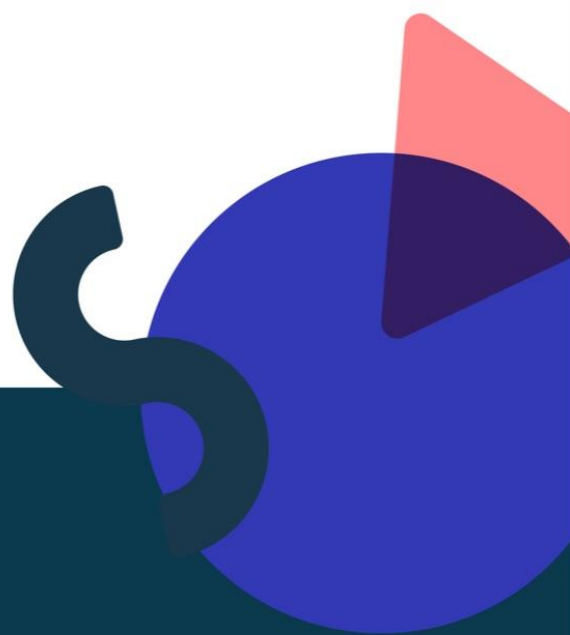
# About SoapBox Labs

Since 2013, SoapBox Labs has been building speech recognition technology that caters to the idiosyncrasies and unique speech patterns of children. Education companies like McGraw Hill, Scholastic, Amplify, and EarlyBird Education use our award-winning technology to voice-enable preK-12 digital screening, practice, intervention, and assessment tools for literacy, math, language learning, and more.

The SoapBox voice engine is proprietary and has been built from the ground up to power immersive and unbiased digital experiences for children ages 2 to 12 of every accent, dialect, and stage of development. SoapBox is a privacy-first and equity-by-design company. In October 2022, SoapBox became the first AI company to be certified for Prioritizing Racial Equity in AI Design by Digital Promise.

Questions? Talk to your SoapBox Account Manager or email us at **hello@soapboxlabs.com**.

# SoapBox