Security of Student Data

Statement on the security of student data and how student PII is being incorporated or fed into the Amira Reading Assistant.

Is Amira a private or public AI Model?

The AI models used by Amira are entirely internally developed, private models. They are closed systems and do not connect to outside sources or 3rd parties for processing. In addition, no student PII is incorporated into Amira's models. Any student data is anonymized before being used to train Amira's AI models.

What AI or LLM is being used?

The AI models used by the Amira product are entirely internally developed, private models. They are closed systems and do not connect to outside sources or 3rd parties for processing.

OpenAI models are used offline, outside of the product to generate some content (e.g. International Phonetic Alphabet pronunciations for words) but any generated content gets reviewed by a human before being used in the product.

Is the AI built within the product or going to outside sources or 3rd parties for processing?

The AI models used by Amira are entirely internally developed, private models. They are closed systems and do not connect to outside sources or 3rd parties for processing.

What types of data are being used to train the AI models?

A subset of student recordings, selected for representativeness across factors such as accent/dialect, age/grade, ELL status, and reading level are used to train Amira's AI models; all student data is anonymized before being included in training data.

What data is being collected from students/ users? Is this data staying within Amira for all processing and AI decision-making?

Student recordings are collected; all data is staying within Amira for all processing and AI decision-making.

How is the data being collected, stored, and secured?

This is outlined in our Privacy Policy, which can be found here:
https://amiralearning.com/amira-privacy

How is data being used for AI training or decision-making?

A subset of anonymized student recordings are labeled by human annotators with respect to whether the student's attempt at each item or word was a correct attempt, and this labeled data is used to train the Amira error detection models.

How is the company addressing potential biases in the AI models?

AI models largely behave based on the composition of the data used to train them. The team takes great care to ensure that the datasets are representative of the population of Amira users. Data sampling is done to ensure populations with different accents, dialects, ages, geographical locations, and reading ability levels are all appropriately represented. The accuracy of the AI models is explicitly measured by the team for subpopulations stratified on the above features.

In addition, DIF analyses have been carried out with respect to factors such as gender and ethnicity by the psychometrics team, based on scoring done by the AI models. Items exhibiting bias based on these analyses on the system's end-to-end results are excluded from the screening.

Are the AI outputs being validated for fairness and or accuracy? If so, how often is this process and output audited?

Every one of Amira's AI models is assessed for both accuracy and fairness at least once a year. Several of Amira's AI models are assessed multiple times a year. Improvements on either accuracy, fairness, or both are shipped via model updates following each assessment.

What processes are in place for testing and validating the AI models for fairness and non-discrimination?

The team explicitly evaluates the accuracy and performance of Amira's models with respect to populations with accent, dialect, ELLs, ages, and geographical locations. These evaluations and model improvements are done at least once a year and are done multiple times a year for some models.

Is there transparency and explainability around the AI models' decision-making processes?

Student recordings along with Amira's scoring of student reading is made completely transparent. Educators have access to student recordings and Amira recordings and can override any scoring they disagree with. Amira has full traceability into the AI models decision-making and upon request can provide detailed data on why a model made the decision it did.

How diverse and inclusive are the teams developing and deploying the AI systems?

Amira's data science team responsible for developing and deploying the AI systems is a diverse and inclusive team. It is a 7-person team led by a woman of color and includes 3 women and 4 persons of color.

What governance and oversight mechanisms are in place for the responsible development and deployment of AI?

There are two primary mechanisms:

1. Amira's AI system architecture is set up with safety as the #1 priority.

- Students do not in any way interact with any external AI models. The AI models used by Amira are entirely internally developed, private models. They are closed systems and do not connect to outside sources.
- Amira's systems are set up such that there is a known, constrained set of things a student can see or interact with. The AI is never generating something that a student will see, it is selecting from a known, constrained set of things that have been created and reviewed by humans.

2. Amira has governance processes in place to ensure safe and responsible AI.

- Any content (e.g. IPA pronunciations for words) generated offline via 3rd party models is reviewed by a human before being used in the platform.
- Amira's software development process for AI models includes a safety review by senior leaders external to the Amira data science team responsible for building the models.
- Every text that goes live in Amira is hand-selected by staff. All suggestions are vetted against our internal rubric. Next, stories are "cleaned", or go through an internal editing process to ensure the story meets Amira's criteria. We

recognize that language is fluid and ever-evolving, so while we do work with content partners whose libraries include stories published during the last decade, we must ensure the stories are updated with best practices such as people-first language, appropriate representation, etc. Finally, the stories are loaded and tested in the product. There are two iterations of testing, pre-and post-publication.

All content that is selected or written for Amira is vetted for:

- Quality metrics
- Appropriate maturity levels for K-6 grade bands
- Objectivity
- Culturally diverse representation
- Student-centered and engaging
- Promotes development of background knowledge
- Promotes vocabulary development
- Decodable *on most scope and sequences

We conduct a quarterly internal quality control review where we run a word analysis on any words that might be questionable in our content library. We review the stories that have the flagged words, and if the story is of any concern we will edit or archive as needed. We also partner with some state agencies and nonprofits who review our content for compliance, quality, and curriculum alignment. Finally, we regularly ask students to review stories and offer feedback on topics, stories, and their experience of content in Amira.